

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 20:53:41 ; Search time 63 Seconds
(without alignments)
8275.405 Million cell updates/sec

Title: US-09-733-685-3
Perfect score: 1700
Sequence: 1 tgggttttattgataaca.....tttacgacttaegaacat 1700

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

- 1: /cggn2_6_ptodata/1/ina/5A_COMB.seq;*
- 2: /cggn2_6_ptodata/1/ina/5B_COMB.seq;*
- 3: /cggn2_6_ptodata/1/ina/6A_COMB.seq;*
- 4: /cggn2_6_ptodata/1/ina/6B_COMB.seq;*
- 5: /cggn2_6_ptodata/1/ina/backfile1.seq;*
- 6: /cggn2_6_ptodata/1/ina/backfile1.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
c 1	103.2	6.1	7218	1	US-08-235-463-14		Sequence 14, Appli
c 2	76.7	4.5	19124	2	US-08-487-26B-13		Sequence 13, Appli
c 3	71.2	4.2	152331	3	US-09-128-155-16		Sequence 16, Appli
c 4	71.2	4.2	176373	3	US-09-128-155-17		Sequence 17, Appli
c 5	62.6	3.7	152331	3	US-09-128-155-16		Sequence 16, Appli
6	58.5	3.4	6124	4	US-08-213-419B-3		Sequence 3, Appli
7	56.8	3.3	289	4	US-09-007-005-17		Sequence 17, Appli
8	56.8	3.3	289	4	US-09-244-796-17		Sequence 17, Appli
9	55.6	3.3	291	1	US-07-922-723A-7		Sequence 7, Appli
10	55.6	3.3	291	1	US-07-793-828G-7		Sequence 7, Appli
11	55.6	3.3	291	1	US-08-074-75-7		Sequence 7, Appli
12	55.6	3.3	291	2	US-08-480-366-7		Sequence 7, Appli
13	55.6	3.3	291	2	US-07-952-277A-7		Sequence 7, Appli
c 14	55.6	3.3	19124	2	US-08-487-826B-13		Sequence 13, Appli
c 15	55.6	3.3	44453	4	US-09-146-106-2		Sequence 5, Appli
c 16	54.2	3.2	5852	1	US-07-867-106-2		Sequence 2, Appli
c 17	53.4	3.1	1956	4	US-08-559-896B-1		Sequence 1, Appli
c 18	52.8	3.1	7379	4	US-09-341-587-5		Sequence 5, Appli
c 19	52.4	3.1	1298	4	US-08-971-089-1		Sequence 1, Appli
20	51.8	3.0	454	2	US-08-622-906A-6		Sequence 6, Appli
21	51.6	3.0	72604	4	US-09-268-992-7		Sequence 7, Appli
22	51.6	3.0	72604	4	US-09-657-474-7		Sequence 7, Appli
23	50.4	3.0	1559	4	US-09-019-095A-7		Sequence 7, Appli
c 24	50.2	3.0	2791	4	US-09-570-367C-1		Sequence 1, Appli
25	50	2.9	1298	3	US-08-948-705-3		Sequence 3, Appli
26	49.6	2.9	417	4	US-08-559-896B-3		Sequence 3, Appli
27	49.6	2.9	1859	3	US-08-691-563C-46		Sequence 46, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DOERNER, F.
; APPLICANT: SCHEIFLINGER, F.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: EP 91 114 300 6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELEPHONE: (703) 636-9300
; TELEX: (703) 633-4109
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: Linear
; IMMEDIATE SOURCE:
; CLONE: PTZSPT-F1S
; US-08-232-463-14
Query Match 6.1%; Score 103.2; DB 1; Length 7218;

Query Match 4.2%; Score 71.2; DB 3; Length 152331;
 Best Local Similarity 54.6%; Pred. No. 1.2e-06;
 Matches 142; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 203 AAATAAAGATGAGATGAGATTCATTGAAAGAAACATGTAGTTG 262
 Db 128751 AGCRAAAGAAAGAAAGAAAGGGAGGAGAAAGAAAGAAAGAA 128692

QY 263 AACAAATAAAGATGAGATGATGATATTTATGAGGTGGTAACATTATTAGG 322
 Db 128691 AGAAAAGAAAGAAAGAAAGAGAGAGGGAGGAGAGAAAGAAAGAG 128632

QY 323 AGGGAGAGAGAAATAGAAAGAAAGAAAGCATGTGAATCTGAAGAAAGATGAAATTGTGTT 382
 Db 128631 AGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 128572

QY 383 AAGATGAGAGAAAGAAAGAAACTCCATGGCTAAAGCTCTGTAAGAGATGAAAGAAAGAA 442
 Db 128571 AAACAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAAAGAAAGAG 128512

QY 443 ACAAAAGAGAAAGAAAGAA 462
 Db 128511 AGGGAGGAGGAAGGAAGAA 128492

RESULT 4
 US-09-128-155-17/C
 Sequence 17, Application US/09128155
 Patent No. 6117654

; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 ; FILE REFERENCE: 09404/052001
 ; CURRENT APPLICATION NUMBER: US/09/128-155
 ; EARLIER APPLICATION NUMBER: US 60/091, 650
 ; CURRENT FILING DATE: 1998-08-03
 ; EARLIER FILING DATE: 1998-07-02
 ; EARLIER APPLICATION NUMBER: US 60/054, 646
 ; EARLIER FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 16
 ; LENGTH: 152331
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE: misc_feature
 ; LOCATION: (1)..(152331)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-128-155-16

Query Match 4.2%; Score 71.2; DB 3; Length 176373;
 Best Local Similarity 54.6%; Pred. No. 1.3e-06;
 Matches 142; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 203 AAATAAAGATGAGATGAGATTCATTGAAAGAAACATGTAGTTG 262
 Db 144714 AGACAAAGAAAGAGAGAGGGAGGAGAAAGAGAAAGAAAGAA 144715

QY 263 AACAAATAAAGATGAGATGATATTTATGAGGTGGTAAGATTATTAGG 322
 Db 144714 AGAAAAGAAAGAAAGAAAGAAAGAGAGAGGGAGGAGAAAGAAAGAG 144655

QY 323 AGGGAGAGAAATGAAAGAAAGAAATGACATGTGAATCTGAAGAAAGATGAAATTGTGTT 382
 Db 144654 AGAAAAGAAAGAAAGAAAGAAAGAGAGAGAAAGAAAGAAAGAG 144595

QY 383 AAAGATGAAAGAGAAAGAAAGAAACTCCATGGCTAAAGCTCTGTAAGAGATGAAAGAAAGAA 442
 Db 144594 AAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAAAGAAAGAG 144535

QY 443 ACAAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 462
 Db 144534 AGGAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 144515

RESULT 5
 US-09-128-155-16
 Sequence 16, Application US/09128155
 Patent No. 6117654

; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 ; FILE REFERENCE: 09404/052001
 ; CURRENT APPLICATION NUMBER: US/09/128-155
 ; CURRENT FILING DATE: 1998-08-03
 ; EARLIER APPLICATION NUMBER: US 60/091, 650
 ; EARLIER FILING DATE: 1998-07-02
 ; EARLIER APPLICATION NUMBER: US 60/054, 646
 ; EARLIER FILING DATE: 1997-08-04
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 16
 ; LENGTH: 152331
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE: misc_feature
 ; LOCATION: (1)..(152331)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-128-155-16

Query Match 3.7%; Score 62.6; DB 3; Length 152331;
 Best Local Similarity 51.28%; Pred. No. 0.0001; Indels 0; Gaps 0;

QY 173 ACATTGAGGAAATTCAGGTTAAAGAGAAATAGAATGAGATGAGATTCT 232
 Db 146914 AGATGGAGGAGGAGGAGAAAGAAAGAAAGAAAGAAAGAAAGAA 146973

QY 233 ATGGAAAAAGAAAGAGAAAGATGTAGTGACAAATAAGAAATATGTGAVATATT 292
 Db 146974 AGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGAAAGAGAAAGAG 147033

QY 293 ATGAGAGGTGTGTGAGATTTTGTAGAGGGAGAGAAAGAAATGAC 352
 Db 147034 AAGGAAGAAAGAAAGAAAGCAAGGAAAGGAAAGGAGGAGGAG 147093

QY 353 ATGGTGAATCTGAAGAGATGAATGTTAAAGATGTTAAAGATGAAAGAGAAAGAGAAACTCCATGG 412
 Db 147094 AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 147153

QY 413 CTAAGCTCTGTTAAAGAGATGAAAGAAACAAAGAAAGAAAGA 457
 Db 147154 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 147198

RESULT 6
 US-08-213-419B-3
 Sequence 3, Application US/08213419B
 Patent No. 633406

; GENERAL INFORMATION:
 ; APPLICANT: Iresburg, J. et al.
 ; TITLE OF INVENTION: GEN Encoding PROTEIN ANTIGENS OF PLASMODIUM FALCIPARUM
 ; FILE REFERENCE: JII-002CNCP
 ; CURRENT APPLICATION NUMBER: US/08/213, 419B
 ; CURRENT FILING DATE: 1994-03-14
 ; PRIOR APPLICATION NUMBER: US 07/870, 506
 ; PRIOR FILING DATE: 1992-04-17
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0

RESULT 9
US-07-922-723A-7
; Sequence 7, Application US/07922723A
; Patent No. 5369004
; GENERAL INFORMATION:
; APPLICANT: Drs. Michael H. Polymeropoulos
; APPLICANT: and Carl R. Merrill
; TITLE OF INVENTION: FIVE HIGHLY INFORMATIVE
; TITLE OF INVENTION: REPEAT POLYMORPHIC DNA MARKERS
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, LeBlanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: DOS Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/922,723A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: D.J. Mills
; REGISTRATION NUMBER: 34506
; REFERENCE/DOCKET NUMBER: 717081B
; TELECOMMUNICATION INFORMATION:
; LENGTH: 291
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-922-723A-7

Query Match 3.3%; Score 55.6; DB 1; Length 291;
Best Local Similarity 52.6%; Pred. No. 0.00065;
Matches 121; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Query 237 AAAAGAAAGAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 296
Db 33 AGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 92
Qy 297 AGAGCTGGTGAAGATTATTAGGGAGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 356
Db 93 AGGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 152

Query Match 3.3%; Score 55.6; DB 1; Length 291;
Best Local Similarity 52.6%; Pred. No. 0.00065;
Matches 121; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Query 357 TGAATCTGAAAGAGATGAATTGTGTAAAGATGAAGAGAAAGAACACTCATGGCTAA 416
Db 33 AGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 92
Qy 417 AGTCCTGTAAGAGATGAAGAGAACAAAATAAGAGATATATGATAATTATTTAG 466
Db 213 AAAGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 262

RESULT 10
US-07-799-828C-7
; Sequence 7, Application US/07799828C
; Patent No. 5378602
; GENERAL INFORMATION:
; APPLICANT: Drs. Carl R. Merrill and
; Michael H. Polymeropoulos
; TITLE OF INVENTION: TWENTY SEVEN HIGHLY INFORMATIVE
; TITLE OF INVENTION: MICROSATELLITE REPEAT
; TITLE OF INVENTION: POLYMORPHIC DNA MARKERS
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, LeBlanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: DOS Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/799,828C
; FILING DATE: 19911127
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: D.J. Mills
; REGISTRATION NUMBER: 34,506
; REFERENCE/DOCKET NUMBER: 717081A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 684 1111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-799-828C-7

Query Match 3.3%; Score 55.6; DB 1; Length 291;
Best Local Similarity 52.6%; Pred. No. 0.00065;
Matches 121; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

Query 237 AAAAGAAAGAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 296
Db 33 AGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 92
Qy 357 TGAATCTGAAAGAGATGAATTGTGTAAAGATGAAGAGAAAGAACACTCATGGCTAA 416
Db 153 AGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 212
Qy 417 AGTCCTGTAAGAGATGAAGAGAACAAAATAAGAGATATATGATAATTATTTAG 466
Db 213 AAAGAAAGAAAGAACATATTAGGGGAGAACAAAATAAGAGATATATGATAATTATTTAG 262

RESULT 11

OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: DOS Text File
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/952,277A
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: D.J. Mills
 REGISTRATION NUMBER: 34506
 REFERENCE/DOCKET NUMBER: 71081C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 703 684 1111
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 291
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-07-952-277A-7

Query Match 3 3%; Score 55.6; DB 2; Length 291;
 Best Local Similarity 52.6%; Pred. No. 0.00065; Matches 121; Conservative 0; Mi smatches 109; Indels 0; Gaps 0;

Qy 237 AAAAGAAAGAGAAACATGTAGTGAAAGAAATAAGAGATGATATTATGGT 296
 Db 33 AGAAAAGAAAAGACCAAGAGACTTGTAGGAGGGAGAGAAATAGACATGG 92

Qy 297 AGAGGTGGTCAAGGTTATTGTAGGAGGGAGAGAAATAGACATGG 356
 Db 93 AGGAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 152

Qy 357 TGAATCTGAGAAAGATGATTGTTAGGAGGGAGAGAAACTCCATGGCTA 416
 Db 153 AGAAGAAAGAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 212

Qy 417 AGTCCTGTAAGAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGA 466
 Db 213 AAAGAAAGAGATGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 262

RESULT 14
 US-08-487-826B-13/C
 ; Sequence 13, Application US/08487826B
 ; Patent No. 5993827
 ; GENERAL INFORMATION:
 ; APPLICANT: Sun, Kim L.
 ; APPLICANT: Chitnis, Chetan
 ; APPLICANT: Miller, Louis H.
 ; APPLICANT: Peterson, David S.
 ; APPLICANT: Su, Xin-Zhuan
 ; APPLICANT: Wellens, Thomas E.
 ; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
 ; NUMBER OF SEQUENCES: 45
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSE: Knobbe Martens Olson & Bear
 ; STREET: 620 Newport Center Drive 16th Floor
 ; CITY: Newport Beach
 ; STATE: California
 ; COUNTRY: US
 ; ZIP: 92660
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/487,826B
 ; FILING DATE: 10-SEP-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned
 REGISTRATION NUMBER: 29,655
 REFERENCE/DOCKET NUMBER: NI1121.001CP1.
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 235-8550
 TELEFAX: (619) 235-0176
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19124 base Pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 US-08-487-826B-13

Query Match 3 3%; Score 55.6; DB 2; Length 19124;
 Best Local Similarity 48.4%; Pred. No. 0.0021; Matches 154; Conservati 0; Mi smatches 164; Indels 0; Gaps 0;

Qy 31 TTATTATTATCATGGTTTATGGATAGCATGAAATAATTATACAGTGTAA 90
 Db 18131 TTTTTTTTTCGCTGGATTATCAGATTAGCATGAAATAACATATATAAA 18072

Qy 91 AACATGTTTGTCTAAAATACATCATTAAATCAGACATTCTTTAAATCAA 150
 Db 18071 CACATATATAAACATTTAAATATAATTATAAAATATCACAAATAATAAA 18012

Qy 151 TCTAACTCTTATATCACAGCACATGACATGAAATTAGGTAAAGAGAAATAAAG 210
 Db 18011 TTAATACTGGAAAAAAATCGAAAATCGAAAATCGAGGATATAATAGAAAAGAG 17952

Qy 211 AATGAGAGATGAGAGAAATTCTATGGAAAGAAAGAGAAACATGAGGAAACATG 270
 Db 17951 AATTATATATATATATATATATATATATATATATAGAAAAATGTACAGAAGAAACAAAT 17892

Qy 271 AAAGAGATGATGATGATATTATGAGGGTGAAGATTATTTAGAGGAGGAGGAG 330
 Db 17891 ACCATATATACATATTAAACAAATAGATACATTAACTAACTAAATATAAA 17832

Qy 331 AGAAATAGAAAAAGAAAA 348
 Db 17831 ATACATATATATAAA 17814

RESULT 15
 US-09-146-053-5
 ; Sequence 5, Application US/09146053A
 ; Patent No. 639349
 ; GENERAL INFORMATION:
 ; APPLICANT: Ravan, James W.
 ; APPLICANT: Sprinkle, Terry Joe Curtis
 ; APPLICANT: Venema, Richard C.
 ; TITLE OF INVENTION: Human Aminopeptidase P Gene
 ; FILE REFERENCE: MCG103
 ; CURRENT APPLICATION NUMBER: US/09/146,053A
 ; CURRENT FILING DATE: 1998-09-02
 ; EARLIER APPLICATION NUMBER: 60/057,854
 ; EARLIER FILING DATE: 1997-09-02
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 44453
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-146-053-5

Query Match 3 3%; Score 55.6; DB 4; Length 44453;
 Best Local Similarity 50.8%; Pred. No. 0.0026; Matches 139; Conservati 0; Mi smatches 139; Indels 0; Gaps 0;

Qy 193 TAAAAAGAGAAAATAAGAAATGAGAGATAGAGAGATTCTATGGAAAAGAGAGAGAA 252

```

Db 20872 TGAAAGGAAAGGAAGAGAAAGAAAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 20931
Qy 253 CATCTAGCGTGAACAAAATAAGAGATATGATGATATTTATGAGGGTGTGAGATT 312
Db 20932 AGAAAGAGAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 20991
Qy 313 ATTTAGGGAGGGAGAGAAATGAAAAAGAAATGACATGGTGAATCTGAAGAAGAT 372
Db 20992 AAGAAAAGAGAGAAAAGGAAAGAAAGAGAAAGAGAAAGAGAAAGAGAAACAA 21051
Qy 373 GAATTGTGTTAACATGAGAGAAAGAGAAACTCCATGGCTAAAGTCTCGTAAGAGAA 432
Db 21052 GAGAAAGAAAGAGAAAGAGAAAGAGAAAGAGAAACAGAGAAAGAGAAAGAAAGAA 21111
Qy 433 TGAAAAGAAACAAAAGAGAGAAAGAAAGAGAAAGAGAAAGGG 470
Db 21112 AAAAGAAAGAAAGAAAGGAGGAGGAGGAGGAGG 21149

```

search completed: November 28, 2002, 23:17:15
 Job time : 816 secs

Result No.	Score	Query	Match Length	DB ID	Description
c 1	76.8	4.5	7657	22 AAS5477	Chemically pretreated base sequence of t
c 2	76.8	4.5	7657	24 ABL54022	Human immune system
c 3	76.8	4.5	14006	24 ABL33404	Chemically treated
c 4	76	4.5	19124	18 AAT72882	Human metastasis a
c 5	76	4.5	19124	21 AA7288287	Human immune system
c 6	74.8	4.4	9539	22 AAS5477	Oligonucleotide fo
c 7	74.8	4.4	9539	24 ABK28180	Oligonucleotide fo
c 8	74	4.4	6644	20 AAX33181	Oligonucleotide fo
c 9	74	4.4	7377	20 AAX33180	Oligonucleotide fo
Run on: November 28, 2002, 20:43:06 ; Search time 268 Seconds (without alignments) 14285.064 Million cell updates/sec	Perfect score: 1700	Sequence: 1 tgggttttattggataaca.....tttagacttacaaat 1700	Scoring table: IDENTITY_NUC Gapop 10.0 , capext 1.0	Minimum DB seq length: 0	Copyright (C) 1993 - 2002 CompuGen Ltd.
Searched: 2185239 seqs, 112599159 residues	Total number of hits satisfying chosen Parameters: 4370478	Post-processing: Minimum Match 0% Maximum Match 100%	Listing first 45 summaries	Database : N_Genesed_101002:*	RESULT 1 ID AAS45477/C XX DT 18-DEC-2001 (first entry)
				1: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1980.DAT;*	DE Chemically pretreated complementary DNA associated with cell cycle #91.
				2: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1981.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				3: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1982.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				4: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1983.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				5: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1984.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				6: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1985.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				7: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1986.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				8: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1987.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				9: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1988.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				10: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1989.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				11: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1990.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				12: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1991.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				13: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1992.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				14: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1993.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				15: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1994.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				16: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1995.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				17: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1996.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				18: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1997.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				19: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1998.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				20: /SIDS2/gcdata/geneseq/geneseq-emb1/NA1999.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				21: /SIDS2/gcdata/geneseq/geneseq-emb1/NA2000.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				22: /SIDS2/gcdata/geneseq/geneseq-emb1/NA2001.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				23: /SIDS2/gcdata/geneseq/geneseq-emb1/NA2001B.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
				24: /SIDS2/gcdata/geneseq/geneseq-emb1/NA2002.DAT;*	XX DE Chemical pretreated complementary DNA associated with cell cycle #91.
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the result being printed.	SUMMARIES				
	8	OS	XX	PN	W0200168911-A2.
		OS	XX	PD	20-SEP-2001.
		OS	XX	PA	(EPIG-) EPIGENOMICS AG.

Db	2568	TATCACTATAACTATAACTTAAACAAAACCCATTACCTTTAACATTCATTCTTC	2509
Qy	145	ATCATTCTAATCTCTTATATCACAGCAGATTGAGGGAAATTCAGGTAAAGAGAAA	204
Db	2508	ATCATTCTAATCAAATAAAAAAAGAAAAAAAGAAAAAAAGAAAAAAAGAAAAAA	2449
Qy	205	ATAAAGAATGAGAGATAGAGATTTCTATGGAAAGAGAACATGTAGGTGAA	264
Db	2448	AAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA	2389
Qy	265	CAAAATAAGAGATGATGATATTTATGAGAGGTGGTAGGATTAGGAG	324
Db	2388	AAAAAAACCAAAATAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAT	2329
Qy	325	GGAGAGAGAAATAGAAAAAAATGACATGGTAATCTGAAGAGATCAATG	384
Db	2328	ATAAAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA	2269
Qy	385	AGATGAAGAGAAAGAGAACTCCATTGGTAAAGTCGTTAAAGAGATGAA	444
Db	2268	AAAAAAATAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA	2209
Qy	445	AAAAGAGGGAGAAAGAGAAAGAGAACCTAAACTATGACTATGGAAAA	496
Db	2208	AAAAAATAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA	2157
RESULT 3			
ABL3958/c			
ID	ABL3958 standard;	DNA; 14006 BP.	
XX	ABL3958;		
AC			
XX	26-MAR-2002	(first entry)	
XX			
DE	Human immune system associated gene	SEQ ID NO: 1931.	
XX	Human; immune system disease; cytosine methylation; antiasthmatic;		
KW	antiarteriosclerotic; antianaemic; cytosstatic; nootropic;		
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;		
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;		
KW	antiinflammatory; arteriosclerosis; anaemia;		
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;		
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene; ds.		
XX	Homo sapiens.		
OS			
XX	WO200200928-A2.		
PN			
XX	03-JAN-2002.		
PD			
XX	PF 02-JUL-2001; 2001WO-EP07537.		
XX	30-JUN-2000; 2000DE-1032529.		
PR			
XX	01-SEP-2000; 2000DE-1043826.		
PA			
PA (EPICG-) EPIGENOMICS AG.			
XX	Olek A, Piepenbrock C, Berlin K;		
XX	WPI; 2002-130909/17.		
DR			
XX	Nucleic acid comprising fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cysteine methylation -		
XX	Claim 1; SEQ ID NO 1931; 32pp + Sequence Listing; German.		
PS	The present invention provides a number of human immune system associated genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory ulcerative bowel diseases. The present sequence is a gene of the invention.		
XX	Sequence 14006 BP; 3278 A; 155 C; 3557 G; 7313 T; 3 other;		
SQ	Query Match 4.5%; Score 76.8; DB 24; Length 14006;		
Best Local Similarity 47.7%; Pred. No. 6.8e-06;			
Matches 225; Conservative 0; Mismatches 247; Indels 0; Gaps 0;			
Qy	25 AAATATTATTTATTCTATGAGTTTATGGATGCCAAATATPATCAT	84	
Db	2628 AAAATAATTAATTTAAACAAACTTTAATACAAAGAACCTTATTCCTATGCC	2569	
Qy	85 GTTAATACATGTTGTTCAATACATGCAATTAAATCAAGATTTAA	144	
RESULT 4			
ID	AAT72882 standard; cDNA; 19124 BP.		
XX			
AC	AAT72882;		
XX			
PD			
XX	DT 12-SEP-1997 (first entry)		
XX			
PF	Plasmodium var-7 gene.		
XX			
DE	DBL gene family; SNBP; sialic acid binding protein; vaccine; therapy;		
XX	Duffy binding like gene; Duffy antigen binding protein; erythrocyte;		
KW	DABP; merozoite; malaria; var-1; var-2; var-3; var-7; immune response;		
KW	Plasmodium; ss.		
XX			
XX	Plasmodium vivax.		
OS			
OS	Plasmodium falciparum.		
XX			
XX	Key Location/Qualifiers		
FH			
FT	7317.15139		
FT	/tag= a		
FT	15140..16205		
FT	/tag= b		
FT	16206..17552		
FT	/tag= c		
FT	/note= "no stop codon given"		
FT			
FT	W0640766-A2.		
PN			
XX	19-DEC-1996.		
PD			
XX	07-JUN-1996;		
PF	96WO-US09508.		
XX			
PR	07-JUN-1995;		
XX	95US-0487826.		
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.,		
XX			
PI	Chitnis C, Miller LH, Peterson DS, Sim KL, Su X;		
PI	Wellens TE;		
XX			
DR	WPI; 1997-052231/05.		
DR	P-PSDB; AAW22475.		
XX			

PT New malaria vaccines - contains cysteine-rich DBL family protein
 PT binding domains homologous domains of the Duffy and sialic acid
 PT binding proteins
 XX
 PS Claim 4; Page 56-61; 96pp; English.
 XX
 CC This sequence represents the var-7 gene of Plasmodium. Var-7 belongs to the Duffy binding like (DBL) family of genes which have homology to the Duffy antigen binding protein (DABP) and sialic acid binding protein (SABP) conserved regions (see AAT7289 and AAT72888 respectively). The var family of genes modulate cytoadherence and antigenic variation of Plasmodium infected erythrocytes. SABP and the Duffy antigen binding protein (DABP), are soluble proteins that appear in the culture supernatant after infected erythrocytes release merozoites. DABP and SABP mediate the binding of merozoites and schizonts to the erythrocyte surface. These proteins are necessary for erythrocyte invasion by the parasite. Thus sequence can be used in the compositions of the invention. The compositions are for the treatment and prevention of malaria, and comprise either a nucleotide sequence or encoded polypeptide of the var-1, var-2, var-3 or var-7 genes of the DBL gene family, a family of genes having homology with conserved regions of DABP and SABP. The compositions are used for the treatment and prevention of malaria. They are also used in the preparation of vaccines for inducing a protective immune response in a mammal to Plasmodium merozoites (especially Plasmodium falciparum or Plasmodium vivax).

XX Sequence 19124 BP: 7824 A; 2190 C; 2790 G; 6320 T; 0 other;

Query Match 4.5%; Score 76; DB 18; Length 19124;
 Best Local Similarity 50.0%; Pred. No. 1e-05;
 Matches 218; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 52 ATGGATAGCAACAAATAATTATATATCAGTGTAAACATGTTCTTAAT 111
 Db 15440 ATTATTAATGAAAAAAGAAAATGAATAATTAAATAATTTAAAAAA 15499
 QY 112 ACATGCATTAAATGACATTTGTTPAATCAATCTCTTATACACAC 171
 Db 15500 AAAAAGAAAAAAAGGAAAAATTTAAATAATTATAATAAANTATA 15559
 QY 172 GACATGAGGAAATTCAGCTAAAGAGAAATAAGATGAGAGATGAGATTTC 231
 Db 15560 AATTGATGATAATAAAAGATATCATCAAAGAAATAAATAAATAA 15616
 QY 232 TATGAAAAAGAACAGAACATGTAGCTAACAAATAAGAGATATGATGATATT 291
 Db 15617 TATATAAAATAATGATTATAAAATAACAAAGAAATAAAACATT 15676
 QY 292 TTATGAGGGTGGTAGAGATTATTTAGGAGGGAGAGAAATAGAAAAAGAAATCA 351
 Db 15677 AAAAAGAAAAAAATATATATATATATAAAACAAAGAAATAATAATAA 15736
 QY 352 CATGGTAATCTGAAAGAGGAATGTAAATGAGAGAAAGAACTCCATG 411
 Db 15737 AAATATATATAAAATAAAATAAAATAAAATAATATATAATAATA 15796
 QY 412 GCTAAAGTCTGTAAGAAGTGAAGAACANAGANGAGAACAGAAAGAGAAAGG 471
 Db 15797 AAATAAAAAAATTTATTAATAATAATAATAATAATAATAATAATAA 15856
 QY 472 TAAATAGACTACTA 487
 Db 15857 AAAAAGATAATAATAATA 15872

QY 292 TTATGAGGGTGGAGAGGAATTTAGGAGGGAGAGAAATGAAAGAAATGAA 351
 Db 15677 AAAAAGAAAAAAATATATCATAAAAACAAAGAAATAATAATAATAA 15736
 QY 352 CATGGTAATCTGAAAGAGGAATGTAAAGATGAGAGAAAGAGAAACTCCATG 411
 Db 15737 AAATATATATAAAATAATAATAAAATAATAATAATAATAATAATA 15796
 QY 412 GCTAAAGTCTGTAAGAAGTGAAGAACANAGANGAGAACAGAAAGAGAAAGG 471
 Db 15797 AAATAAAAAAATTTATTAATAATAATAATAATAATAATAATAATAA 15856
 QY 472 TAAATAGACTACTA 487
 Db 15857 AAAAAGATAATAATAATA 15872

DE plasmodium var-7 polypeptide encoding DNA.
 XX DBL gene; Duffy-binding like gene, ebl-1; Duffy Antigen Binding Protein;
 KW DABP; Sialic Acid Binding Protein; SABP; malaria; vaccine; immunisation;
 KW protozoa;cide; var-7; ds.
 XX Plasmodium sp.
 PN US5993827-A.
 XX PD 30-NOV-1999.
 XX PF 07-JUN-1995; 95US-0487826.
 XX PR 10-SEP-1993; 93US-0119677.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI Sim RL, Chitnis C, Peterson DS, Su X, Wellens TE, Miller LH;
 XX DR WPI; 2000-194198/17.
 DR P-PSDB; AAY77904.
 XX Isolated protein binding domains from Plasmodium vivax and Plasmodium falciparum erythrocyte binding proteins useful for vaccinating against malaria -
 XX Disclosure; Columns 91-108; 93pp; English.
 PS
 XX The invention relates to ebl-1 polypeptides that are encoded by the DBL (Duffy-binding like) gene family. The ebl-1 proteins are substantially identical to the Duffy Antigen Binding Protein (DABP) and Sialic Acid Binding Protein (SABP), which are soluble proteins that appear in the culture supernatant after erythrocytes infected with malaria release merozoites. Immunological studies indicate that DABP and SABP are the respective ligands for Plasmodium falciparum Duffy and sialic acid receptors on erythrocytes. The ebl-1 polypeptides may be used to vaccinate against malaria, especially caused by P. falciparum. Immunization with the polypeptide provides effective protection against malaria. The present sequence represents the var-7 polypeptide encoding DNA.
 Sequence 19124 BP; 7824 A; 2190 C; 2790 G; 6320 T; 0 other;
 Query Match 4.5%; Score 76; DB 21; Length 19124;
 Best Local Similarity 50.0%; Pred. No. 1e-05;
 Matches 218; Conservative 0; Mismatches 215; Indels 3; Gaps 1;
 Db 15440 ATTATTAATGAAAAAAGAAAATGAATAATTAAATAATTTAAAAAA 15499
 QY 112 ACATGCATTAAATGACATTTGTTPAATCAATCTCTTATACACAC 171
 Db 15500 AAAAAGAAAAAAAGGAAAAATTTAAATAATTATAATAAANTATA 15559
 QY 172 GACATGAGGAAATTCAGCTAAAGAGAAATAAGATGAGAGATGAGATTTC 231
 Db 15560 AATTGATGATAATAAAAGATATCATCAAAGAAATAAATAAATAA 15616
 QY 232 TATGAAAAAGAACAGAACATGTAGCTAACAAATAAGAGATATGATGATATT 291
 Db 15617 TATATAAAATAATGATTATAAAATAACAAAGAAATAAAACATT 15676
 QY 292 TTATGAGGGTGGTAGAGATTATTTAGGAGGGAGAGAAATAGAAAAAGAAATGAA 351
 Db 15677 AAAAAGAAAAAAATATATCATAAAAACAAAGAAATAATAATAATAA 15736
 QY 352 CATGGTAATCTGAAAGAGGAATGTAAATGAGAGAAAGAACTCCATG 411
 Db 15737 AAATATATATAAAATAAAATAAAATAATATATAATAATAATA 15796
 QY 412 GCTAAAGTCTGTAAGAAGTGAAGAACANAGANGAGAACAGAAAGAGAAAGG 471
 Db 15797 AAATAAAAAAATTTATTAATAATAATAATAATAATAATAATAA 15856
 QY 472 TAAATAGACTACTA 487
 Db 15857 AAAAAGATAATAATAATA 15872

RESULT 5
 ID AAZ98287 standard; DNA: 19124 BP.
 XX AC AAZ98287;
 XX DT 13-JUN-2000 (first entry)
 XX

Db 15677 AAAAAGAAAAAAATATATCATAAAAACAAAGAAATAATAATAATAA 15736
 QY 352 CATGGTAATCTGAAAGAGGAATGTAAAGATGAGAACTCCATG 411
 Db 15737 AAATATATATAAAATAAAATAAAATAATATATAATAATAATA 15796
 QY 412 GCTAAAGTCTGTAAGAAGTGAAGAACANAGANGAGAACAGAAAGAGAAAGG 471
 Db 15797 AAATAAAAAAATTTATTAATAATAATAATAATAATAATAA 15856
 QY 472 GCTAAAGTCTGTAAGAAGTGAAGAACANAGANGAGAACAGAAAGAGAAAGG 471

QY	67	AAATATTAAATATATCAGTGTATAAACATGTTCTAAATACATGGCTTAAATTAATTAAATAAA 15856	
Db	589	AAAAAAAATTTAATTTAAATAAAAAATATAAAATATAAAATATAAAATATAAA 530	
QY	127	TCAAGATTTCTTTAAATCAAATCTATCAGCTATCACAGCATGACGAAA 186	
Db	529	AAAAAAAATTCATCAGCTATCACAGCATGACGAAA 470	
QY	187	TTCAGTTAAAGAGAAATAGAGATAGAGATTCTATGGAAAAAAAG 246	
Db	469	AAAAAAAATTTAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 410	
QY	247	AGAGAACATCTAGGTGACAACAAANTAAAGAGATATGATGATATTTATGAGAGTTGTG 306	
Db	409	AAAAAAAACGAAAAAAACAAAAAAACAAAAAAATAAAAAAATAAAAAA 350	
QY	307	AAGATTTTATGGAGGGAGAGAATAGAAAAGAAAATGATGGTAATCTGAA 366	
Db	349	AAAAAAAATTTAAATAAAAAAACAAATAAAATAAAAAAACAAA 290	
QY	367	GAAGATGAAATTGTGTTAAAGTGAAGAGAAAGAACCTCATGGCTAAAGTCTCGTAA 426	
Db	289	AAAAAACAAAGCACAAAAAAACGAAAAAAACGAAAAAAATGACACAAAA 230	
QY	427	AGAAGATGAAAAAGAACACAAAGAGGAAAGAGAAAGGCTTAAATAGACTAATCT 486	
Db	229	ATAAAACGAAAAATACAAAAAAATACAAAAAAATACAAAAAAATACAAAAAA 170	
QY	487	ATTGCCAAAA 496	
Db	169	AAAAAAA 160	
		RESULT 7	
		ABK28180/c	
		ID ABK28180 standard; DNA: 9539 BP.	
		XX	
		AC ABK28180;	
		XX DT 23-APR-2002 (first entry)	
		DE DNA transcription associated complementary genomic DNA #27.	
		XX	
		KW DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;	
		KW PNA; cytosine methylation state; sup; retroviral infection; gene; ds;	
		KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;	
		KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;	
		KW immunological disorder; Werner syndrome; developmental disorder;	
		KW psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;	
		KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;	
		KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;	
		KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;	
		KW polyglutamine disorder; solid tumour.	
		XX Unidentified.	
		XX WO200192565-A2.	
		XX PD 06-DEC-2001.	
		XX PF 06-APR-2001; 2001WO-EP03973.	
		XX PR 06-APR-2000; 2000DE-1019058.	
		PR 07-APR-2000; 2000DE-1019173.	
		PR 30-JUN-2000; 2000DE-103529.	
		PR 01-SEP-2000; 2000DE-1043826.	
		XX PA (EPIG-) EPIGENOMICS AG.	
Olek A,		Piepenbrock C, Berlin K;	
WPI; 2001-602751/68.			
X		Designing primers and probes for analysing diseases associated with	
X		cytose methylation state e.g. arthritis, cancer, aging,	
X		arteriosclerosis comprising fragments of chemically modified genes	
X		associated with cell cycle -	
X		Claim 1: SEQ ID No 52: 28pp; English.	
X		Sequences AAS45296-AAS4520 represent chemically pretreated genomic DNA	
C		molecules associated with the cell cycle and specific PCR primers of the	
C		invention. The sequences are useful for detecting the methylation state	
C		of all CpG dinucleotides in a sequence and therefore for analysing	
C		associated diseases. By analysing cytose methylation patterns in the pretreated	
C		DNA, genetic and/or epigenetic parameters for the diagnosis and therapy	
C		of existing diseases or the predisposition to specific diseases can be	
C		ascertained. The parameters may be compared to another set of genetic	
C		and/or epigenetic parameters, the differences serving as basis for	
C		diagnosis and/or prognosis events which are disadvantageous to patients.	
C		The sequences of the invention are useful for the diagnosis and therapy	
C		of HIV infection, neurodegenerative disorders, graft-versus-host disease,	
C		aging, glomerular disease, Lewy body disease, arthritis,	
C		arteriosclerosis, solid tumours and cancers.	
X		Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 other;	
X		Query Match 4.4%; Score 74.8%; DB 22; Length 9539;	
Q		Best Local Similarity 48.4%; Pred. No. 1.6e-05;	
Q		Matches 208; Conservative 0; Mismatches 222; Indels 0; Gaps 0;	
Q		DR WPI; 2002-090046/12.	
Q		PA (EPIG-) EPIGENOMICS AG.	
Q		Olek A, Piepenbrock C, Berlin K;	

PT New nucleic acids or oligomers, useful for diagnosing or treating diseases associated with DNA transcription, e.g., immunological disorders, Werner syndrome, psoriasis, myocardial infarction, solid tumours or cancer

XX

PS Claim 1; SEQ ID No 54; 32pp; English.

XX The invention relates to a nucleic acid, which comprises a segment of the chemically pretreated DNA of genes associated with DNA transcription from one of 346 sequences, and an oligomer, in particular, an oligonucleotide or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical to the chemically pretreated DNA of genes associated with DNA transcription. The set of oligomer probes are useful for detecting the cytosine methylation state and/or single nucleotide polymorphisms (SNPs) in a chemically pretreated genomic DNA. The nucleic acids are useful for diagnosing or treating diseases associated with DNA transcription (particularly with the methylation status), e.g. adenosine deaminase deficiency, viral infection, retroviral infection, Sezary syndrome, haematological disorders, immunological disorders, Werner syndrome, tuberculosis, developmental disorders, psoriasis, Rieger's syndrome, neurological disorders, neurodegenerative disorders, Waardenburg syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial infarction, hypertension, angiogenesis, erythropoiesis, congenital heart disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours or cancer. Sequences ABK28127-ABK28472 represent DNA transcription associated genomic DNA molecules of the invention.

CC Note: The sequence data for this Patent did not form part of the printed specification but was obtained in electronic format directly from the European Patent Office.

XX Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 other;

SQ Query Match 4.4%; Score 74.8; DB 24; Length 9539;

Best Local Similarity 48.4%; Pred. No. 1.6e-05;

Matches 208; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 67 AAATTTAATATAACATGTTAAATACATGCATTAAAAA 126

DB 589 AAAAAGAAAATAAACAAAACAAAAAAACAAAAAAA 330

QY 127 TCAGACATTGTTAAATCAAATCTTAATTCATCACACGACATTGCGGGAAA 186

DB 529 AAAAAGAAAATCAAAACAAAAAAACAAAAAAACAAAAAA 470

QY 187 TTCAAGTAAAGAGAGAAAATAAGATGAGATAAGAGATTCATGGAAAAGAG 246

DB 469 TAAGAAAATATAAAAGAAAATAAAAGAAAATAAAATAAAA 410

QY 247 AGAGAACATGTAGCTGACAAATAAGAGATATGATGATATTTATGAGGGTGGTG 306

DB 409 AAAAAGAACGAAAAAAACAAAAAAACAAAAAAATAAAA 350

QY 307 AGAGTATTAGGAGGGAGAGAAATGAGAAAAATGACATGGTAAATCGAA 366

DB 349 AAAAAGAACGAAAAAAATAAAACAAATAAAATAAAAACAAAAAAATAAAA 290

QY 367 GAGAGGAATGTGTTAAATGAGAGAACATGGTAAATCTCGTAA 426

DB 289 AAAAAGAACGAAAAAAACAAAAAGAACGAAAGAACGAAAGAACGAAAC 230

QY 427 AGAAGATGAAAGAACGAAAGAACGAAAGAACGAAAGAACGAAAGAACGAAACT 486

DB 229 ATAAAACGAAATAATCACAAAATAAAACAAAAAAACAAAAAA 170

QY 487 ATTGCCAAA 496

DB 169 AAAAAGAAA 160

PT	AC	AAX33181;
PT	XX	25-JUN-1999 (first entry)
PT	DE	Base sequence of the plasmid pRx-ires-bsr.
XX	KW	Cowpox virus; bsr; viral vector; expression; apoptosis; resistance; crma; bcl-2; bcl-xL; FLIP; survivin; IAP; ILP; adenovirus; cancer; autoimmune disease; graft rejection reaction; inflammation; inflammatory disease; ss.
PS	XX	Synthetic.
XX	OS	Cowpox virus.
XX	PN	WO9113073-A2.
XX	PD	18-MAR-1999.
XX	PF	07-SEP-1998; 98WO-JP04010.
XX	PR	08-SEP-1997; 97JP-0259235.
XX	PA	(RPR-) RPR GENCELL ASIA PACIFIC INC.
XX	PI	Hamada H;
XX	DR	WPI; 1999-243728/20.
XX	PT	New apoptosis-resistant virus-sensitive cell
XX	PS	Example 1; Page 38-41; 51pp; English.
XX	CC	The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, tie treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant gene introduced is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRx-ires-bsr, which contains the cowpox virus bsr gene, and is used in an example from the present invention.
SQ	CC	Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 other;
QY	CC	Query Match 4.4%; Score 74; DB 20; Length 6644;
DB	CC	Best Local Similarity 48.4%; Pred. No. 2.1e-05;
DB	CC	Matches 206; Conservative 0; Mismatches 220; Indels 0; Gaps 0;
QY	QY	107 AAAATACATGCATTAAATCAGACATTGTTAAATCAAATCTCTTATATC 166
DB	DB	4018 AAAAAGAAAATATAAAAGAAAATAAAATAAAATAAAATAAA 4077
QY	QY	167 ACAACGACATTGAGGAAATTCAAGTAAATAGAGATAGAGATAGAG 226
DB	DB	4078 AAAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 4137
QY	QY	227 ATTCTCATGGAAAAGAGAACATGAGGTAAATAAGAGATATGATGAT 286
DB	DB	4138 AAAAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 4197
XX	QY	287 ATATTTATGAGGGTGGTAAAGATTTTAGGAGGGAGAGAAATGAGAAAT 346

RESULT 8
AAX33181 standard; DNA; 6644 BP.
ID AAX33181 XX

WPI; 1999-243728/20.		XX	DE Base sequence of the plasmid pRx-Bcl 2-i-hcd 25.
New apoptosis-resistant virus-sensitive cell		XX	XX Cpoox virus; bsr; viral vector; expression; apoptosis; resistance; crmA; bcl-2; bcl-xL; FLIP; survivin; IAP; ILP; adeno-virus; cancer; autoimmune disease; graft rejection reaction; ss; inflammatory disease; ss.
Example 1; Page 34-38; 51pp; English.		XX	The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis, by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the cowpox virus bsr gene which is used in an example from the present invention.
Sequence 7797 BP; 2542 A; 1760 C; 1656 G; 1839 T; 0 other;		XX	Sequence 7797 BP; 2542 A; 1760 C; 1656 G; 1839 T; 0 other;
Query Match 4.4%; Score 74; DB 20; Length 7797;		PS Example 3; Page 46-49; 51pp; English.	PS Example 3; Page 46-49; 51pp; English.
Best Local Similarity 4.84%; Pred. No. 2.2e-05;		XX	XX The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis, by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRx-Bcl 2-i-hcd 25, which contains the human Bcl-2 gene, and is used in an example from the present invention.
Matches 206; Conservancy 220; Indels 0; Gaps 0;		XX	XX
107 AAAATACATGCTATTAAATCAGACATTGTTPAACATCAATCTATCTTATATC 166	b	PT New apoptosis-resistant virus-sensitive cell	PT New apoptosis-resistant virus-sensitive cell
5171 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5230	b	XX	XX
167 ACAAAGCACATTGACGGAAAATTCAAGGTTAAAGAGAAAATAAGAATGAGAGAG 226	y	PS Example 3; Page 46-49; 51pp; English.	PS Example 3; Page 46-49; 51pp; English.
5231 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5290	b	XX	XX
227 ATTCTCTATGGAAAAAGAAGAGAACATGTAGGTGAAACAAAATAAGAGATATGATGAT 286	y	CC The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis, by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRx-Bcl 2-i-hcd 25, which contains the human Bcl-2 gene, and is used in an example from the present invention.	CC The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis, by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRx-Bcl 2-i-hcd 25, which contains the human Bcl-2 gene, and is used in an example from the present invention.
5291 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5350	b	CC	CC
287 ATATTTATGAGGTGTGAGATTATTGTAGGAGGGAGAACATAAGAAAGAGAA 346	y	CC	CC
5351 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5410	b	CC	CC
347 ATAGACATGGTAATCTGAAGAGATGAATTGTGTAAAGATGAAGAGAAACT 406	y	CC	CC
5411 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5470	b	CC	CC
407 CCTGGCTAAAGTCGTAAGAAGATGAAAGAAGAACAAAAGAGAAAGAGA 466	y	CC	CC
5471 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5530	b	CC	CC
467 AAGGCTAAATAGACTAATCTGCCTAAATTCTGTAGCGAACAACTATTTGGTC 526	y	CC	CC
5531 AAAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5590	b	CC	CC
527 AAGGTT 532	y	Db 5370 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5429	Db 5370 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5429
5591 AAGCTT 5596	b	Db 5430 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5489	Db 5430 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5489
RESULT 11		QY 167 ACAAGCACATTGACGGAAAATTCAAGCATTTGTGTTAAATCAATCTCTTATTC 166	QY 167 ACAAGCACATTGACGGAAAATTCAAGCATTTGTGTTAAATCAATCTCTTATTC 166
AAX33184 standard; DNA; 7996 BP.		Db 5490 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5549	Db 5490 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5549
AAX33184;		QY 227 ATTCTCTATGGAAAAAGAAGAGAACATGTAGGTGAAACAAAATAAGAGATATGATG 286	QY 227 ATTCTCTATGGAAAAAGAAGAGAACATGTAGGTGAAACAAAATAAGAGATATGATG 286
D		Db 5550 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5609	Db 5550 AAA.....AAAAAAA.....AAAAAAA.....AAAAAAA.....AAAAAAA 5609

QY	347	ANTGACATGGTGAATCTGAAAGAAGATGAAATGTGTAAAGATGAGAGAAAGAGACT	406	QY	TTTGTGTTCTTAAATAACATGCCATTAAATCAGACATTTGGTTTAAAA[CAATCTATACT	157					
Db	5610	AAAAAAA[A[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[5669		Db	51850	TCTPATCATTAATTTCACACTAAACTCCPATAACTTAAACAA[AAAAAAATCTTC	51791				
Qy	407	CCATGCCTAAAGTCGTAAGAAGATGAAAGAAGAAACAAGAAAGAAGAAGAGA	466	Qy	158	TCTPATCATCACACGATTGAGAAATTCAGGTA[AGAGAAATAAAAGATGAGA	217				
Db	5670	AAAAAAA[A[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[5729		Db	51790	TA[AAAAACCAACATAAACATTACACCATAAAATATAA[AAAAAA[AAAAAA[51731					
Qy	467	AAGGCTAAATAGACTAACTATGGCAAATTTCUTGTAGCCGACAAATACTATTGGGCC	526	Qy	248	GATAGAGAGATTCTATGAAAAGAAGAGAGAAAC[GTAGGTGACAAATAAGAGA	277				
Db	5730	AAAAAAA[A[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[5789		Db	51730	AAAAAA[AAAAACATTACRADA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[51671					
Qy	527	AAGGTT	532	Qy	278	TATGATGATATAATTAGAGGGTAGGATTATTAGAGAG[GAGAGAAATA	337				
Db	5790	AAGCTT	5795	Db	51670	AAAAAA[AAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[51611					
<hr/>											
RESULT 1.2											
ABL34174/C				QY	338	AAAAAA[AAAAATGACATGGTGAATCTGAGAGATGATTTGGTTAAAGATGAGAGAGA	397				
ID	ABL34174	standard; DNA; 113515 BP.		Db	51610	AAAAAA[AAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[51551					
XX				Qy	398	AAGAGACTCCATGGCTAACAGATGCTAAAGAGATGAAAAGAACAAAGAAGAGA	457				
AC	ABL34174;			Db	51550	AAAAAA[AAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[AAAAAA[51491					
XX				QY	458	AGRAAGAGAAAGGCTAAATAGACTAACTATGCCAAATTC	500				
DT	26-MAR-2002	(first entry)		Db	51490	AAAAAA[AAAAACACAAAA[AAAAAA[AAAAAA[AAAAAA[51448					
DE	Human immune system associated gene	SEQ ID NO: 2147.		RESULT 1.3							
XX				ABL33404/C							
KW	Human; immune system disease; cytosine methylation; antiasthmatic;			ID	ABL33404 standard; DNA; 16033 BP.						
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;			XX	ABL33404;						
KW	antiinflammatory; antiarthritic; antidiabetic; antipsoriatic;			XX	26-MAR-2002 (first entry)						
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;			DE	Human immune system associated gene SEQ ID NO: 1377.						
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;			XX	XX						
gene; ds.				XX	XX						
OS	Homo sapiens.			XX	XX						
XX				XX	XX						
PN	WO200200938-A2.			XX	XX						
XX				XX	XX						
XX				PD	03-JAN-2002.						
XX				PF	02-JUL-2001; 2001WO-EP07537.						
XX				PR	30-JUN-2000; 2000DE-1032529.						
XX				PR	01-SEP-2000; 2000DE-1043826.						
PA	(EPIG-) EPIGENOMICS AG.			XX	Olek A, Piepenbrock C, Berlin K;						
XX				XX	WPI; 2002-130909/17.						
PI	Olek A, Piepenbrock C,	Berlin K;		XX	PN	WC200200928-A2.					
XX				XX	XX						
DR				PD	03-JAN-2002.						
XX				PF	02-JUL-2001; 2001WO-EP07537.						
XX				PP	02-JUL-2001; 2001WO-EP07537.						
PS	SEQ ID NO 2147; 32bp + Sequence Listing; German.			XX	XX						
XX				PR	30-JUN-2000; 2000DE-103229.						
PT	Nucleic acid comprising fragment of chemically modified gene, useful			PR	01-SEP-2000; 2000DE-1043826.						
PT	for diagnosis and treatment of diseases associated with abnormal			XX	(EPIG-) EPIGENOMICS AG.						
PT	cytose methylation -			XX	Olek A, Piepenbrock C, Berlin K;						
XX				XX	WPI; 2002-130909/17.						
PS	Query Match 4.3%; Score 73.4; DB 24; Length 113515;			XX	XX						
PS	Best Local Similarity 48.9%; Pred. No. 4.8e-05;			XX	Claim 1; SEQ ID NO 1377; 32pp + Sequence Listing; German.						
XX	Mismatches 0; Matches 197; Conservative 0;			XX	The present invention provides a number of human immune system associated						
CC	Sequence 113515 BP; 31803 A; 1174 C; 24020 G; 56518 T; 0 other;			CC	genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel diseases. The present sequence is a gene of the invention.						

GenCore version 5.1.3
copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 28, 2002, 21:46:16 ; Search time 3105 seconds
 (15933.903 Million cell updates/sec)

Title: US-09-733-685-3

Perfect score: 1700

Sequence: 1 tgggttttatggataaca.....tttacgacttacgaaatat 1700

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:
 1: gb_ba:
 2: gb_htg:
 3: gb_in:
 4: qb_on:
 5: qb_ov:
 6: qb_pa:
 7: qb_ph:
 8: qb_pi:
 9: qb_pr:
 10: qb_ro:
 11: qb_sts:
 12: qb_sy:
 13: qb_un:
 14: qb_vt:
 15: em_ba:
 16: em_fnt:
 17: em_hum:
 18: em_in:
 19: em_mu:
 20: em_om:
 21: em_or:
 22: em_ov:
 23: em_pat:
 24: em_ph:
 25: em_p1:
 26: em_ro:
 27: em_sts:
 28: em_vt:
 29: em_vt:
 30: em_htg_hum:
 31: em_htg_inv:
 32: em_htg_other:
 33: em_htg_mus:
 34: em_htg_p1n:
 35: em_htg_rod:
 36: em_htg_mam:
 37: em_htg_rt:
 38: em_sy:
 39: em_htgo_hum:
 40: em_htgo_mus:
 41: em_htgo_other:

Result No.	Score	Query	Match	Length	DB	ID	Description
c 1	1700	100.0	83650	8	AB023041		AB023041, Arabidopsis
c 2	576.2	33.9	92620	8	AB026636		AB026636, Arabidopsis
c 3	540.2	31.8	83646	8	AB005248		AB005248, Arabidopsis
c 4	534.6	31.4	94487	8	AC012394		AC012394, Arabidopsis
c 5	534.6	31.4	100806	8	AC015450		AC015450, Arabidopsis
c 6	499.8	29.4	104386	8	ATT3A17		ATT3A17, Arabidopsis
c 7	499.8	29.4	179771	8	ATCRIV25		ATCRIV25, Arabidopsis
c 8	497	29.2	95519	8	AF071527		AF071527, Arabidopsis
c 9	497	29.2	116448	8	AC005142		AC005142, Arabidopsis
c 10	497	29.2	159629	8	ATCRIV9		ATCRIV9, Arabidopsis
c 11	369	21.7	369	8	ATH50185		ATH50185, Arabidopsis
c 12	330.8	19.5	95190	8	AC007203		AC007203, Arabidopsis
c 13	103.2	6.1	7218	6	T6694		T6694, Sequence 14
c 14	90.2	5.3	171763	2	AC15878		AC15878, Mus muscu
c 15	89.2	5.2	280982	2	AC125184		AC125184, Mus muscu
c 16	87.4	5.1	161362	2	AC101939		AC101939, Mus muscu
c 17	87.4	5.1	217779	2	AL844515		AL844515, Mus muscu
c 18	86.8	5.1	202872	2	AC016160		AC016160, Homo sapi
c 19	86.6	5.1	166299	2	AC17244		AC17244, Mus muscu
c 20	85.6	5.0	44735	9	AC02315		AC02315, Homo sapi
c 21	85.6	5.0	48623	2	AC110469		AC110469, Homo sapi
c 22	85	5.0	178038	10	AI672308		AI672308, Mouse DNA
c 23	84.6	5.0	161362	2	PDON04		PDON04, P.domestica
c 24	84.6	5.0	61953	2	AC101966		AC101966, Mus muscu
c 25	84.6	5.0	177648	10	AL732328		AL732328, Mouse DNA
c 26	84.6	5.0	228944	2	AL844530		AL844530, Mus muscu
c 27	84.4	5.0	222556	2	AC126558		AC126558, Mus muscu
c 28	84	4.9	404763	5	AF044763		AF044763, Cecropia
c 29	84	4.9	202756	2	AC119804		AC119804, Mus muscu
c 30	84	4.9	218893	2	AC03839		AC03839, Mus muscu
c 31	83.6	4.9	162803	2	AC117818		AC117818, Mus muscu
c 32	83.6	4.9	177648	10	AC102446		AC102446, Mus muscu
c 33	83.4	4.9	239851	2	AC101712		AC101712, Mus muscu
c 34	83.4	4.9	247340	2	AC121591		AC121591, Mus muscu
c 35	83.2	4.9	64789	2	AC083839		AC083839, Mus muscu
c 36	82.8	4.9	167364	10	AC121590		AC121590, Mus muscu
c 37	82.8	4.9	220257	10	AL732240		AL732240, Mouse DNA
c 38	82.6	4.9	48659	2	AC115677		AC115677, Dictyoste
c 39	82.4	4.8	178413	2	AL805914		AL805914, Mus muscu
c 40	82.4	4.8	203295	10	AL663049		AL663049, Mouse DNA
c 41	82.2	4.8	204839	2	AC113318		AC113318, Mus muscu
c 42	82	4.8	163132	2	AC097180		AC097180, Rattus no
c 43	81.6	4.8	93661	9	AC107463		AC107463, Homo sapi
c 44	81.6	4.8	106815	10	AL645809		AL645809, Mouse DNA
c 45	81.6	4.8	120635	2	AC127463		AC127463, Ornithorh

SUMMARIES

ALIGNMENTS

RESULT 1
 AB023041/c

LOCUS AB023041

DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, p1 clone: MPE11.

ACCESSION AB023041 BA000014

VERSION AB023041.1 GI:4220640

KEYWORDS SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui p1 clone:MPE11.

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis; 1 (sites)

REFERENCE Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., and Tabata, S.

Pred. No. is the number of results predicted by chance to have a

TITLE	Structural analysis of <i>Arabidopsis thaliana</i> chromosome 3. I. Sequence features of the regions of 4,504,864 bp covered by sixty p1 and TAC clones	CDS
JOURNAL	DNA Res. 7 (2), 131-135 (2000)	
MEDLINE	20277480	
AUTHORS	Sato, S., Nakamura, Y., Kaneko, T., Kato, T., Asamizu, E. and Tabata, S.	
JOURNAL	Submitted (01-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-5, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)	
COMMENT	Address for correspondence: kato@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?=&MPEI1 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K922 and the 3' clone is MJL14.	
FEATURES	Location/Qualifiers	
source	/organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="Taxon:3702"	
exon	/number=1 /clone_id="MPEI1_5" /clone_lip="Mitsui F1" /complement="542..764") /gene_id:K922_5"	
CDS	/evidence="not_experimental join(1705..243,258..2739,2862..2949,3037..3107, 3190..3213,3410..3514) /note="gb:AAD55139.1 /gene_id:MPEI1_1" /codon_start=1 /product="dihydroxyacetone phosphate acetyltransferase" /protein_id="BAB01047.1" /db_xref="GI:9279589" /translation="MTVRSPREIFYDQSYLAATVPGGETAPGAGIGLIAETEEAKSKAASKSSSS ESDKADMDFYEDYDQSYLAATVPGGETAPGAGIGLIAETEEAKSKAASKSSSS VAEAVVPPSPPVTPSSPAPAAQPAEVTAVSGPKTATPYAKKLQHKVDIESTAG TGPFGRITASDVETAGIAFSKSSAAPPPPPVTAKTATNPLPQDPSIVPTTA MOSAVAKNMESLSVPTFQYGPNTDQPKTATPYAKKLQHKVDIESTAG PVNASCCKDKSKSNSNSINAVAVAINGLIPVLOQADKDDYLUSQKWKELVKA RSKQLOPHENGTPTLSNQGMFQYDREFALLPQGQAIMAVGASKPTVADKGFFS VKNTMLVNVADHRIVYGADLAAFLQTFAKTTENPDSTL"	
CDS	/complement(4594..5106) /note="gene_id:MPEI1_2 /unknown_protein" /codon_start=1 /evidence="not_experimental /protein_id="BAB01048.1" /db_xref="GI:9279590" /translation="MEDLIERFIRLPTDVSQNKRVRGDLIDSDPDKVRLDDFDGQL DVSQQLD5WKSFEINELSTTAALSSGQTQDUGLYFEASDDELGFIPPTPQTL PPSCETETVTLVRASSDSSVEGLCGFEDHVTFFGCPDGLGQFLFYFDGCCDSDGL	

	TRLLVYFLIRFPESSGRSSWSGTGRRSRRPASLNGESYEGNTGT SPRDEENNOSSDEOVSGBTRRSRRTVDDDEEEEL" * complement(join(32197, . . . 32430, 32525. . . 32676, 32969. . . 33332, 33404. . . 34487)) /note="gene_id: MPE11_9" /codon_start=1 /evidence-not_experimental /product="chloroplast 50S ribosomal protein L15" /protein_id="BAB01055_1" /db_xref="GI: 927959;" /translation="MATPSIISNNPLTSSRHCYRLHSSUSEKGNVSLGANPSQISSL KLNQNLTKTRNQOQPARIIVVQSTAAATCAGCAAAATACATTATGGCTAGCC ISAGOGSGGMRQSKSSPGIMRGFGGQPALEYRIPKURGIAQMRSGLPKLYP VNIKDLTEAQFOEGLDEVSETLTKOKGLINPSGRERKLPKLUKGELSMKLFKEKAF STQAREKLEESGCPLTVLGCRKWKPSAKRNDAADEFFAKRAAARAEARTSEPARS A" CDS complement(join(33961. . . 34021, 34114. . . 34226, 34391. . . 34482, 34602. . . 34800)) /note="gb:AAF26483_1 gene_id:MPE11_10 similar to unknown protein" /codon_start=1 /evidence-not_experimental /protein_id="BAB01056_1" /db_xref="GI: 927958;" /translation="MKNYMLJIDNESENASYDLIWALENQDTESSKVYIFAKOPONS FTPPPNVLSSEVQIIFYFSPNSELNRILQAENMKIILIGLEKKKICLNLIGIKAE FTDGDPKDLRKIQRNINLIVTSQDSLKCTQNIDCSLLVKKRKLD" join(36033. . . 36150, 36243. . . 36283, 36497. . . 36651. . . 36707, 36797. . . 36871) /note="contains similarity to transcription-associated zinc ribbon protein gene_id:MPE11_11" /codon_start=1 /evidence-not_experimental /protein_id="BAB01057_1" /db_xref="GI: 927959;" /translation="MEKSRESEFLFCNLCGTMVLKSYAECPHCKTRNARDIIDK EIATVNSAETIRELGIUSFEKTOQAEALPKKKKACERPCQHELEVTTQFSADES QTYYTCPCNAHRTEG" join(37285. . . 37965, 38579. . . 38674)	Query Match 100.0%; Score 1700; DB 8; Length 83650; Best Local Similarity 100.0%; Pred. No. 2.1e-285; Matches 1700; Conservative 0; Mis.matches 0; Indels 0; Gaps 0;
Qy 1	TGGGTGTTAATGGATAACATGACAATAATTGTTATTCATGAGTTTATGGATAG Db 12680 TGGGTGTTAATGGATAACATGACAATAATTGTTATTCATGAGTTTATGGATAG	60 12621
Qy 61	CATGACAATAATTATAATCATGTTAATAACATGTTTCTAAATAACATGCATT Db 12620 CATGACAATAATTATAATCATGTTAATAACATGTTTCTAAATAACATGCATT	120 12561
Qy 121	TAAAATCAGACATTGTTAATCAAACTCAATCTCTTATCACAAAGACATTGAC Db 12560 TAAAATCAGACATTGTTAATCAAACTCAATCTCTTATCACAAAGACATTGAC	180 12501
Qy 181	GCAAATTCAGGTAAGGAAATAAGGAAATAAGGATAGAGATTCATGGAAA Db 12500 GCAAATTCAGGTAAGGAAATAAGGAAATAAGGATAGAGATTCATGGAAA	240 12441
Qy 241	AGAAAGAGAGAACATGTTGAAACAAATAAGGATAGATGATATTTATGAGAC Db 12440 AGAAAGAGAGAACATGTTGAAACAAATAAGGATAGATGATATTTATGAGAC	300 12381
Qy 301	CTGGTGAAAGATTATTTAGGAGGGAGAGAGAAATAAGGAAATAAGGATTCATGGTGA Db 12380 CTGGTGAAAGATTATTTAGGAGGGAGAGAGAAATAAGGAAATAAGGATTCATGGTGA	360 12321
Qy 361	TCTGAAGAAAGATGTTAAAGATGAGGAGAAAGAGAAACTCCATGGCTAAAGTC Db 12320 TCTGAAGAAAGATGTTAAAGATGAGGAGAAAGAGAAACTCCATGGCTAAAGTC	420 12261

QY	1501	GGCTTGTCTGGCCGTTAGTAATGCAAGTAGGTATGCCGTATAAGGGT	1560
Db	11180	GGCTTGTCTGGCCGTTAGTAATGCAAGTAGGTATGCCGTATAAGGGT	11121
QY	1561	CCAAAAAAGACGGCATTCGGTTGGCTTGGAACTTGGATTGGATTAGT	1620
Db	11120	CCAAAAAAGACGGCATTCGGTTGGCTTGGAACTTGGATTGGATTAGT	11061
QY	1621	CATGGTTATCTATTAAATGCTCGGGACTTGTGGACGAGCGCTTGCTCTCTG	1680
Db	11060	CATGGTTATCTATTAAATGCTCGGGACCTGTGGACGAGCGCTTGCTCTCTG	11001
QY	1681	TTPACGACTTACGAACATAT	1700
Db	11000	TTPACGACTTACGAACATAT	10981

RESULT	2						
AB026636	AB026636	92620	bp	DNA	linear	PLN 27-DEC-2000	
LOCUS	Arabidopsis thaliana	genomic	DNA,	chromosome	3,	TAC	clone:K14A17.
DEFINITION	ACCESSION	AB026636	BA000014				
VERSION	AB026636.1	GI:4757392					
KEYWORDS							
SOURCE	Arabidopsis thaliana (strain:columbia)	DNA,	clone_lib:Mitsui TAC				
ORGANISM	Arabidopsis thaliana						
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;							
Spermatophyta; Magnoliophyta; eudicots; core eudicots;							
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis;							
1 (sites)							
REFERENCE	Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E. and Tabata, S.						
AUTHORS							
TITLE	Structural analysis of Arabidopsis thaliana chromosome 3. I.						

SEQUENCE LECTURES OR THE REGIONS OF 4,504,004 bp COVERED BY SIXTY	
JOURNAL	PLANT AND TAC CLONES
MEDLINE	DNA RES. 7 (2), 131-135 (2000)
REFERENCE	20277450
AUTHORS	SATO, S., NAKAMURA, Y., KANEKO, T., KATO, T., ASAMIZU, E. and TABATA, S.
TITLE	Direct Submission
JOURNAL	Submitted (28-APR-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd.cgi?=&K1A17 Genes with similarity to proteins in the database are described as product, or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
	The software programs used to predict genes include: Graill

(Informatics Group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/Graill-1.3/>).
 GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>),
 NetGene2 (S.M. Rebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>), and
 SplicerPredictor (Volker Brendel, Stanford University,
<http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi>).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submission.
 The 5' clone is MUH15 and the 3' clone is MCE21.

source
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3102"

```
/chromosome="3"  
/clone="K14A17"  
/clone_1lib="Mitsui TAC  
complement(201:2180)
```

58

TTVRNHDISLHDINTLWDDELLVPTDNNQSAPTFVNDSYTFEPLQNAANGTAEHGNA
 THSNSNIPPEOENITUDSYSPAWAYSEEKTTKLVLIGNSFLHUSYOHLRSNLNYCVGCDCEVPA
 EYLQACVYRCITPPHSQWNNLTLSDAHCKPISQCFCREHRAVPVLDTKTPBDDNQSK
 WEEFEROVRSHLLEFTSSNRLVNSKPNLRAKLAESTNHLNSNAYLVTSIQ
 GNKYFDQADHLEPLSLSKRNLMKVKLVEMLNTTYDSKGICGHLSLGLTWS
 VOLFSISLISLNFRDKQGHTALWHAAYGREKWAALLSAGARPNLVTSDPKNDLUGC
 MADDIAQNSYDGLMKAQFQKAAKSMLAGNTGIDKEAKMNLNOCTLPEDQ
 SLKDLAQATTAEEARLIOQAFEEKAQKAAKSMLAGNTGIDKEAKMNLNOCTLPEDQ
 RKYDERPKIEFAAYRQLQCHOTWTKIRRELYNMRQAIQARAFGLQRQTKLILWSV
 GVLERAVLWRQKRGFLGRQLQVAQEEADEFTKTSQQAEEERESVRYQAM
 .FRSKEAQDQYTRMLTHAQPQVNLTHLTFNLNSPKKNMR
 complement(join(18170, .18172, 18407, .18490, 18736, .18807,
 18977, .19090, 19168, .19281, 19387, .19487, 19590, .19668,
 19755, .19805, 19906, .19989, 20035, .20431, 20539, .20669,
 20916, .21040, 21126, .21231, 21349, .21483, 21624, .21770,
 21860, .22171))
 /note="gene_id : K14A17_6"
 /pseudo
 /codon_start=1
 /evidence="not_experimental
 /product="lipopamide dehydrogenase"
 join(29056, .29205, 29695, .29826, 29925, .30009, 30149, .3031267)
 /note="contains similarity to similar to ubiquitin
 conjugating enzyme
 gene_id : K14A17_7"
 /codon_start=1
 /evidence="not_experimental
 /protein_id="BAA94978_1"
 /db_xref="GI: 7670024"
 /translation="MADERYNRKNPVAKRILQEVKEMQANPSDFMSPLPLEENIFEQ
 FAIRPGDIDFEGGYIHYRQLADYTFKPPQFMLLTQNGFETNTKICLISYHPE
 HWQPSWSWRTALVIALAMPMTSPNGAISGSDYPKDERRTLAKSRETPTYQSPRQK
 LIDEHQJILSKLAVVPPRLPLPCSQASIVTSEVYKASVSEPOEAIVVEEISLATDTI
 VDDQDIEATAEYNTAAVSVPAPEVYKASVSEPOEAIVVEEISLATDTI
 WAAYGLTAVMULLKKTKNSYSTCFMDDS"
 /note="#g01A0D26962_1
 join(32260, .32365, 33121, .33412, 33506, .33750, 33837, .3411
 gene_id : K14A17_8
 similar to unknown protein"
 /codon_start=1
 /evidence="not_experimental
 /protein_id="BAA94979_1"
 /db_xref="GI: 7670025"
 /translation="MGISQSNVYDRIKEROMLSFFKIFQARDISSLDCMRALPYSEIS
 KVSRKDFSYKVMVRAQNCKTWDVEVSXKPNPTVYIETRGWDFLVSQDNEPGR
 RGNRNFHNNYIYONGLERMKPKRFQTMGSPSSKIKKEGENSLIDYKKEESEDEPGR
 EFLYRKTTKEDSSKSKMTRUKKKSXSKVQVLGDVFEKFTIRKSYLKELAIPK
 HFDYDHINKSKLFTIRHPNGGSWKVLLCILVREIRTFSGGSKLAREFPLMVGKCTF
 KLIKPFEVLLISKKNRKEKMDQCM"
 complement(join(34528, .34725, 35213, .35322, 35680, .35863,)
 /note="#g01A0F26101_1
 gene_id : K14A17_9
 similar to unknown protein"
 /codon_start=1
 /evidence="not_experimental
 /protein_id="BAA94980_1"
 /db_xref="GI: 7670026"
 /translation="MASSGGRIGYAVDFSDCSKKALSWAIDNVYRDGIDLILITIAH
 DMNYYEGGMQMLWETGSPFIMSEFSDAYMKVALKPKDAETLDOVNTARKKTITVV
 MKLYWGDREKICAAAEQIPPLSLVNGRIGGLKRMIGSVSNHVNNAACP7VVK
 AHI"
 complement(join(36309, .36416, 36556, .36846, 37250, .37324,
 37500, .37654, 38020, .38145, 38221, .38374, 38433, .38573,
 38878, .38987, 39134, .39281, 39474, .39551, 39993, .40562,)
 /codon_start=1
 /evidence="not_experimental
 /protein_id="BAA94981_1"
 /db_xref="GI: 7670027"
 /translation="MGDTNGASLIELGQDQEVEOPFLARDYARTVISPEDDEKEE"

CDS	join(42817, . 43766, 43839, . 43979, 44083, . 44403) /note="emb CA886928.1 gene_id:MXI10_10 similar to unknown protein" /pseudo	Db 23160 /codon_start=1 /evidence=not_experimental join(45089, . 45677, 46843, . 47158, 47312, . 47452, 47544, . 47871) /note="contains similarity to non-LTR retroelement reverse transcriptase gene_id:MXI10_11" /pseudo /codon_start=1 /evidence=not_experimental 48978, . 49433 /note="gene_id:MXI10_12 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09152.1" /db_xref="GI:758838" /transl_id="MHWISMKRVSAYVEGKERNNSNGEETRREGSKHARMTAKAR MNVIKESMRPLPPKRKKSMIISLMRKEDQDNVTVKEKEKGTVONIRE NAEMSSKVEENHMLYEEENTWFAFQNGWQRYDVMNSGKSNEN" complement(join(50633, 50872, 51043, . 51177, 51268, . 51438)) /note="gene_id:MXI10_13 sp P10798" /codon_start=1 /evidence=not_experimental	QY 601 TTTTCCACACAT 613 Db 23172 TATTAAATAT 23160
CDS	RESULT 4 AC012394 LOCUS Arabidopsis thaliana chromosome 1 BAC F15M4 genomic sequence, PLN 12-OCT-2000 DEFINITION complete sequence.	QY 94487 bp DNA F15M4 genomic sequence, PLN 12-OCT-2000 Db 23172 TATTAAATAT 23160	
CDS	AC012394 AC012394 AC012394.3 GI:6554469 HTG. SOURCE Arabidopsis thaliana. ORGANISM Arabidopsis thaliana. Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. REFERENCE 1 (bases 1 to 94487) AUTHORS Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, J.H., Haas, B.J., Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M. TITLE Arabidopsis thaliana chromosome 1 BAC F15M4 genomic sequence JOURNAL unpublished REPERCUE 2 (bases 1 to 94487) AUTHORS Lin, X. and Kaul, S. TITLE Direct Submission JOURNAL Submitted (27-OCT-1999) The Institute for Genomic Research, 9712 REPERCUE 3 (bases 1 to 94487)	AC012394 AC012394 AC012394.3 GI:6554469 HTG. SOURCE Arabidopsis thaliana. ORGANISM Arabidopsis thaliana. Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. REFERENCE 1 (bases 1 to 94487) AUTHORS Lin, X., Kaul, S., Town, C.D., Benito, M.-I., Creasy, J.H., Haas, B.J., Wu, D., Maiti, R., Ronning, C.M., Koo, H., Fujii, C.Y., Utterback, T.R., Barnstead, M.E., Bowman, C.L., White, O., Nierman, W.C. and Fraser, C.M. TITLE Arabidopsis thaliana chromosome 1 BAC F15M4 genomic sequence JOURNAL unpublished REPERCUE 2 (bases 1 to 94487) AUTHORS Lin, X. and Kaul, S. TITLE Direct Submission JOURNAL Submitted (27-OCT-1999) The Institute for Genomic Research, 9712 REPERCUE 3 (bases 1 to 94487)	
CDS	Query Match 31.8%; Score 540.2; DB: 8; Length 83646; Best Local Similarity 93.6%; Pred. No. 2e-84; Title: Gaps 1; Matches 574; Conservative 0; Mismatches 38; Indels 1; Gaps 1; /evidence=not_experimental	QY 1 TGGGTTTATGGATAACATGACAATATTATTTATGGTTTATGGATAG 60 Db 23771 TGGGTTTATGGATAACATGACAATATTATTTATGGATAG 23712 QY 61 CATGACAATTAATATACAGTTAAACATCTATCATGTTCTAAATACATGCATT 120 Db 23711 CATGACAATTAATATACAGTTAAACATCTATCATGTTCTAAATACATGCATT 23652 QY 121 TAAATTCAGACATTGTGTTAAATCAAACTCATCTATCACAGACATGAC 180 Db 23651 TAAATTCAGACATTGTGTTAAATCAAACTCATCTATCACAGACATGAC 23592 QY 181 GGAAATTCAAGGTAAAAGAAAAAAGAATGAGAGATAGAGATTTCTATGGAAA 240 Db 23591 GGAGATTCAAGGTAAAATAAAATTAAGAATGAGAGATGGTTCTATGGAAA 23533 QY 241 AGAAAGAGAACATGTAGGTGACAAATAAGAGATGAGTATATTATGAG 300 Db 23532 AGAAAGAGAACATGTGGGTGACAAATAAGAGATGAGTATATTATGAG 23473 QY 301 GTGGTGAAGATTATTAGGAGAGGAGAGAAATAGAACATGGTGAAG 360 Db 23472 GTGGTGAAGATTATTAGGAGAGGAGAAATAGAACATGGTGAAG 23413 QY 361 TCTGAAGAACATGAAATTCTGTAGCCGACAAACTATGGCTAAAGTC 420 Db 23412 TCTGAAGAACATGAAATTCTGTAGCCGACAAACTATGGCTAAAGTC 23353 QY 421 TCCTAAAGAGATGAAAAAGAACAAAAAGAGAAAGAGAAAGCTAAATAGA 480 Db 23352 TCCTAAAGAACATGAAATTCTGTAGCCGACAAACTATGGCTAAAGTC 23293 QY 481 CTAACATTGCGCAAATTTCTGTAGCCGACAAACTATGGCTAAAGTC 540 Db 23292 CTAAATTAATGCCAACATTCTGTAGCCGACAAACTATGGCTAAAGTC 23233 QY 541 TATCTTCTTCAGTCATAATGTTCTCATATACTCTAAATATAGCCGATACAA 600 Db 23232 TATCTTCTTCAGTCATAATGTTCTCATATACTCTAAATATAGCCGATACAA 23173	QY 94487 bp DNA F15M4 genomic sequence, PLN 12-OCT-2000 Db 23172 TATTAAATAT 23160 AC012394 AC012394 AC012394.3 GI:6554469 HTG. SOURCE Arabidopsis thaliana chromosome 1 BAC F15M4 genomic sequence, PLN 12-OCT-2000 ORGANISM Arabidopsis thaliana. Eukaryota; Viridiplanteae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; Eudicotyledons; Core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. REFERENCE 1 (bases 1 to 94487) AUTHORS Lin, X., Kaul, S. TITLE Direct Submission JOURNAL Submitted (12-OCT-2000) The Institute for Genomic Research, 9712 REPERCUE 5 (bases 1 to 94487) AUTHORS Lin, X. and Kaul, S. TITLE Direct Submission JOURNAL Submitted (05-OCT-2000) The Institute for Genomic Research, 9712 REPERCUE 6 (bases 1 to 94487) AUTHORS Lin, X. and Kaul, S. TITLE Direct Submission JOURNAL Submitted (12-OCT-2000) The Institute for Genomic Research, 9712 REPERCUE 7 (bases 1 to 94487) AUTHORS Lin, X. and Kaul, S. TITLE Direct Submission JOURNAL On Dec 10, 1999 this sequence version replaced gi:6143858. COMMENT Address all correspondence to:at@tigr.org BAC Clone F15M4 is from Arabidopsis thaliana chromosome 1. The orientation of the sequence is from SPC to T7 end of the BAC clone. Genes were identified by a combination of several methods: Gene prediction programs including Genscan+ (Chris Burge, http://ccr-081.mit.edu/GENSCAN.html), GeneMarkHMM (Mark Borodovsky, http://genemark.biology.gatech.edu/GeneMark/), GlimmerM (a variant of GlimmerM, see Mihaela Peretea, http://www.tigr.org/software/glimmer_hun/glimmer.html), and GeneSplicer (Mihaela Peretea and Steven Salzberg, contact mperetea@tigr.org), searches of the complete sequence against a peptide database and the plant EST database at TIGR (http://www.tigr.org/tgi.shtml). Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as hypothetical proteins. Genes encoding tRNAs are predicted by

/gene="F14G6.3"	mRNA	complement(join(<26257 . 26401,26478 . >27736), 27062 . 27247,27332 . 27486,27593 . >27736))
/codon_start=1		/gene="F14G6.7"
/product="putative phosphate transporter; 18176-15417"		/note="contains nonconsensus splice site"
/protein_id="AAG51941.1"		/complement(join(26257 . 26401,26478 . >27736), 27062 . 27247,27332 . 27486,27593 . >27736))
/db_xref="GI:12323970"		/gene="F14G6.7"
/translation="MPEELSLSLDAARIOWHFKALIVAGMGLFTDAYDLCFCIAPIM KMTSQIYHRDSITGALLTSYALIUTGALQQLIGFLGDRVKVKGULLINVF SSPCCGFSYCTTRESCVAVSLLGFFRFLGCGDYLSATISMEFANKRGRGIAFAA VFSNOGLCILMSASSANTVOTCAKNAARGSSKPTKNTVAGLTLAPPESDIARLILMIG ALPAFLTFWRLMPETARYTALENVNVQAAMDQRMMSVMSIQTEDSSELEQP PSSSYKUFSRREPLSHGRDLFAASANWLFVFTSNLQSIQINFNSPQNLSINV WWSXHEKINGFLGVFLWTSYACSTPGVWFLWTSYACSTPGVWFLWTSYACSTPGVWFL YDSAAFEVAKLAATVAACTGFLWTSYACSTPGVWFLWTSYACSTPGVWFLWTSYACSTPGVWFL LVEAEACTGFLWTSYACSTPGVWFLWTSYACSTPGVWFLWTSYACSTPGVWFL KEVRPLSAEKLKLNGLWKFRPLEETIDSVSVEAAGDLKVA"	/note="contains nonconsensus splice site"	
/codon_start=1		/product="putative cinnamoyl-CoA reductase; 27707-26257"
/protein_id="AAC51951.1"		/db_xref="GI:12323980"
/translational="MAVKORVCVGTEGFFIASLVKFLLSRGYTVHGTVRDPCEDEKND HLRKLDNASKNKLKFADLFEDDEGFSADGCSCHVHASPVPFEGEHLKPALTGK NLVLEACTETVKQKVUVVSSAAVYNPKWQDVAKDEDWSDFQYLHSERYYYLAK LTERALEWSKRNFADVVTCPSVLTGPRQSTLQHPLKIGGIKSLUSDELY LVDVDRDADALLYVNEATGRYTCNSHSLYTSMLMEKLKNYPKRNPESFTEVKE KEVRPLSAEKLKLNGLWKFRPLEETIDSVSVEAAGDLKVA"	/note="similar to cinnamoyl-CoA reductase GB:CAA56103 (Bucalyphtus gunnii); contains non-consensus GG acceptor splice site at exon 4."	
complement(<18673 . 18926,19124 . >19301))	Query	31.4%
/gene="F14G6.4"	Match	Score 534 6; DB 8;
/note="predicted by genscan"	Best	Length 100806;
complement(<18673 . 18926,19124 . >19301))	Local	Pred. No. 1.8e-83;
/gene="F14G6.4"	Matches	Gaps 1;
/codon_start=1	Conservative	Indels 1;
/product="hypothetical protein; 19301-18673"	0;	Mismatches 19;
/protein_id="AAC51943.1"		
/db_xref="GI:12323972"		
/translation="MNAENNOTTTTHSKVSHVCTGTAKLGSGVPPIGLYDIGVSEY AYIFRVSLPIEKKDKICIEQREGRCVQIYGVPETAPSITGCLYRMRQVQLCPFG PFSTIFNLPQDPRLFSPNFRSDGLFVEVVVKLGVRIPTS"		
AYIFRVSLPIEKKDKICIEQREGRCVQIYGVPETAPSITGCLYRMRQVQLCPFG PFSTIFNLPQDPRLFSPNFRSDGLFVEVVVKLGVRIPTS"	QY	1 TGGGTTTTATGGATACTGACAATATTATTCATGAGTTTATTGGATAG 60
20920 . 22411	Db	39846 TGGGTTTTATGGATACTGACAATATTATTCATGAGTTTATTGGATAG 39787
/gene="F14G6.5"	QY	61 CATGACAATATTATTCATGAGTTTAAATCAAACTCATCTTATATCACACGACATTGAC 120
/product="hypothetical protein; 21472 . 21330,21407 .>2308,"	Db	39786 CATGACAATATTATTCATGAGTTTAAATCAAACTCATCTTATATCACACGACATTGAC 39727
join(<20920 . 20986,21073 . 21148,21227 .>2308, 21139 . 21281,21911 . 21986,22064 . 22182,22264 . 22308, 22385 . 22411)	QY	121 TAAAATCAGACATTGCTTTAAATCAAACTCATCTTATATCACACGACATTGAC 180
/gene="F14G6.5"	Db	39726 TAAAATCAGACATTGCTTTAAATCAAACTCATCTTATATCACACGACATTGAC 39667
/product="hypothetical protein; 21472 . 21330,21407 .>2308,"	QY	181 GAAAATTCAGTTAAAGAGAAATAAGAGATGAGAGATCTATGGAAA 240
join(<20920 . 20986,21073 . 21148,21227 .>2308, 21139 . 21281,21911 . 21986,22064 . 22182,22264 . 22308, 22385 . 22411)	Db	39666 GGAAATTCAGGTAAAGAGAAATAAGAGATGAGAGATGAGATCTATGGAAA 39607
/gene="F14G6.5"	QY	241 AGAAAGAGAACTATGGTGTGAC-ANAAATANGAGATGAGATATTTATGAGA 299
/codon_start=1	Db	39606 AGAAAGAGAACTATGGTGTGACAAACAAATAAGAGATGAGATATTTATGAGA 39547
/product="unknown protein; 20920-22411"		
/protein_id="AAC51945.1"		
/db_xref="GI:12323974"		
/translation="MAAISWLSSPOSESNPYRPTITDSRRCSSISAATSVLDSNEEQ HRISSDHVGKRRDMLQIASSYFLPLIAISPAFAETNASEAERVYTDENTNKPEISI PQVSTAITGCPDFTRMESGKVEAFTLVSLDLSRWSKPGVTAKLIDSRSATKGFY FI"		
complement(<23340 . 23431,23548 . 23785,23879 .>24808))	QY	300 GGTGCTGAGATTTTATGGAGGGAGAGAAATAAGAAAGAAA:GACATGGTGA 359
/gene="F14G6.6"	Db	39546 GGTGCTGAGATTTTATGGAGGGAGAGAAATAAGAAAGAAA:GACATGGTGA 39487
/codon_start=1	QY	360 ATCTGAAAGAACTATGGTGTAAAGTGAGAGAAAGAAACTCATGGCTAAAGT 419
/product="Unknown RNA recognition motif. (k.a. RRM, RBD, or RNP domain)"	Db	39486 ATCTGAAAGAACTATGGTGTAAAGTGAGAGAAAGAAACTCATGGCTAAAGT 39427
/protein_id="AAC51945.1"		
/complement(<23340 . 23431,23548 . 23785,23879 .>24808)	QY	420 CTCTAAAGAACTATGGTGTAAAGTGAGAGAAAGAAACTCATGGCTAAAGT 479
24008 . 2414 . 24234 . 24378,24602 .>24808))	Db	39426 CTCTAAAGAACTATGGTGTAAAGTGAGAGAAAGAAACTCATGGCTAAAGT 39367
/gene="F14G6.6"	QY	480 ACTRACTATTGCCAAAATTTCTGAGCACATACTATGGTCACAGGTATTTGT 539
/complement(<23340 . 23431,23548 . 23785,23879 .>24808))	Db	39366 ACTAAATATTGCCAAAATTTCTGAGCACACTATGGTCACAGGTATTTGT 39307
/gene="F14G6.6"	QY	540 GTATTCCTTTGAAGTCAAAAGTTTCTTACATATA 576
/codon_start=1	Db	39306 GTATTCCTTTCAACTCAAAGTTTCTTACATATA 39270
/product="putative RNA-binding protein; 24808-23340"		
/protein_id="AAC51948.1"		
/db_xref="GI:12323977"		
/translation="MAYQPVPGSGFHYLNSPFDGTTFTKVFVGLAWETQSETLRQHFEYGEILEAVVYIADKNTGRSGYGYFTFRACPFEEAARPGDPLPILQPMPELLEMCRVLEPHLEIRISHHFHTTSKELSLIMVTA GPBMYSQSQQGYSYSPMGQQLQYVGPVAVNSPYQYGGLSQFIPNGIGYTAVQGYS VPGSHLQLQGGPTVSTMITSMPALQAPYPSGIPGPAPVSHLTVHSPOMSTASDQ TIR"		
complement(<26257 . 27736))	RESULT	6
/gene="F14G6.7"	ATT32A17/c	LOCUS
/note="similar to cinnamoyl-CoA reductase GB:CAA56103 (Bucalyphtus gunnii); contains non-consensus GG acceptor splice site at exon 4."	Arabidopsis thaliana DNA chromosome 4, BAC clone r32A17 (ESSA project)	
	AL161813 AC007635	
	AL161813.1 GI:321058	
	thale cress.	

ORGANISM *Arabidopsis thaliana*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1
 (bases 1 to 96353)
AUTHORS McCombie, R.W., Spiegel, L.A., Huang, E.N., Nascimeto, L.U., de la Bastide, M., VIL, D.M., Preston, R.R., Matero, A., Shah, R., O'Shaughnessy, A., Rodriguez, M., Shekher, M., Schutz, K., See, L.H., Swaby, I., Habermann, K., Dednia, N.N., Mewes, H.W., Lemcke, K. and Mayer, K.F.X.

JOURNAL Unpublished
 2 (bases 95732 to 104386)
AUTHORS McCombie, R.W., Robben, J., Gromponez, B., Bastiaens, I., Volckaert, G., Mewes, H.W., Lemcke, K. and Mayer, K.F.X.

REFERENCE EU Arabidopsis sequencing project.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz, 8a, D-8213 Martinsried, FRG, E-mail: leimcke@mips.biochem.mpg.de, Mayer@mips.biochem.mpg.de
COMMENT On Mar 23, 2001, this sequence version replaced gi:5103797. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.mpg.de/proj/thal/>.

FEATURES

source	Location/Qualifiers
	1..104386
	/organism="Arabidopsis thaliana"
	/variety="Columbia"
	/db_xref="taxon:3702"
	/chromosome="4"
	1..22908
	/note="overlap to BAC T3F12; Please refer to EMBL:AC002983 or analysis and annotation"
	22865..24293
	/gene="AT4g08700"
	complement(join(22865..23096,23227..24293))
	/gene="AT4g08700"
	complement(join(22865..23096,23227..24293))
	/note="similarity to various hypothetical proteins of Arabidopsis thaliana"
	Contains prokaryotic membrane lipoprotein lipid attachment site AA82-92; Prokaryotic membrane lipoprotein lipid attachment site AA95-135
	contains EST gb:AT597220.1"
	/codon_start=1
	/product="putative protein"
	/protein_id="CAB82106..1"
	/db_xref="GI:1321059"
	/translation="MILLKEEDEGSTRISVPTOLMKLKTRHWHWLYFISIFFLISAQATAVLIGRFYNEGGNSKIVSTLQYQTCGPILYPLCFLPASHISSSCFKYLWYIYL
	SLGPAIGUNLILYSGKLVLSSAYSTSILCQLSQAFANGVSYTINSQKICLCLLSEVLF
	LSVAVLYSLDDNSPSGDSKWSLIGCLCVTFASLYSLQSLSMOFSEFNVLKSET
	FSWYLEMQYTSLVASCYAVIGLIFASGLWMLSVEMEEFHCGQIVTYFLTGAVSWLG
	LSGVAVALFLYSLSFNLIGPLSLVYKPLAIAVHDKJLNEVKAYAMLAFMGIGCP
	YIXONYLDDLKVQBARAQAEYLFLRAKVLCLYIRNSTEVKINGFIENLGDPHV
	VSEHDVAELTNWESNHVDOSTLKAASSKRKMVKFGF"
	complement(22865..23096)
	/gene="AT4g08700"
	/number=1
	complement(23097..23226)
	/gene="AT4g08700"
	/number=1
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"
	/number=2
	complement(23227..24293)
	/gene="AT4g08700"
	/number=2
	complement(23097..23226)
	/gene="AT4g08700"

/gene="At4g08820"
 /note="similarity to various hypothetical proteins of
 Arabidopsis thaliana Chromosome 11, III and IV"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB1007_1"
 /db_xref="GI:7267525"
 /translation="MSVSNKGSVGGGTLKPKENVIDDAFVVERVLLEFFPENGDFGPVIT
 ISPEVLLFAMGLWKRCKMVKFPDGFVGYLWQDWSPRF
 PLRDIIVTIPWVRLTPNVPNYHRCLEELARGLKLKYDLNTTFRGRARVCI
 EVNLAKPKLPGTLYNGDRYFAEVEEGFTVYRQSGIRVEKPDKMVTAQSSGRS
 NRSURELFKNQVEISNRGGLEEDWASGEVALLSNKTKGKMRKESVT
 QVEAQVNVTFKKGKGENGGPKWKGNGLTIMEPNSPROQGSYNKPVQLGLFGPTK
 GEESNNVYDALSANGKLRVVEQRSGRPGGAFVSGLETSGEMDSMELPONMNQIQ
 SKFLICDATAVY"
 exon
 8902..9093
 /gene="At4g08820"
 /number=1
 9094..9192
 /gene="At4g08820"
 /number=1
 exon
 9193..9553
 /gene="At4g08820"
 /number=2
 9554..1685
 /gene="At4g08820"
 /number=2
 exon
 9686..10347
 /gene="At4g08820"
 /number=3
 10891..14348
 /gene="At4g08830"
 join(10891..11025,11244..11393,11430..12120,12361..13089,
 13162..13585,13634..13633)
 /gene="At4g08830"
 /note="similarity to various putative reverse
 transcripts of Arabidopsis thaliana
 Contains Aminoacyl-tRNA synthetases class-II
 signatures AA699..708"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB70008_1"
 /db_xref="GI:7267526"
 /translation="MAPDSLEFGENINRVS1IDMGFSGNKYTWKBGRVENTFYAKRLD
 RLKKWNNDVFGDVKKKEELAREIQVQDALDVHSDMDMERKEELIKEFDKSREKWI
 ALGDGRINNFHTTIRRQNVRERLHNSTGWTDAKEIHALAQYRLLYSNDVWL
 QVNPFLPDQGIERLTHBEVLSNKPLATEVEAVERSMDKYYKAQPDGTQPIFQSCW
 EVVGNSVIRFALDFESSGLPPOTNDALWVLPKPKPMNQRE-TSLCNAIFKMIT
 KMMVLRLIKVIGKLIGPSQSFIPGRSLSDNIVVQEAEHSLNRRKGGERLCHIMDR
 AAVAKWIKSISQGOKPSKTCFADLLUFAEAWSORVIRRLTECTIASGOKVS
 LDKSKEIFSKVNSRDLIEKLISKESEKGLKSPTELGKLYGMPLORLNUKDFGEVLERVS
 SRLAGKKGSSLSSAGLELTFSLIPMTSILPOSTLEGIDKLERVFLGSSA
 EKKKLHVAWRVCLPKSEGLGIRTSKCNKMKLVSKVWRLLINDRYSWARLRSKY
 RVGLREYVSRSRWYVNGRDILFEDSNWNSHEALINRAVIEINSEKELRVIDLWAN
 GLGNKLDKTEPYTSWHLLELAIVCFDLSVARDLSWGTSADEGVFTVSAYRJLTD
 HDPRPNNAAFDRLRVAVALERVKFLWHIGDTSYCQVKGGDETILHFLKDPSIAG
 IWRRLYVQSYDFTGSGLFWGLWLYNLGKMAENGYANTLETAIVVWNWISKWRCGYF
 GEVGRKDRKVKFPERDIAETSHAHHSQNGGLKTRVERLVAWKPDDGWVKINTDGA
 SRGNLGLATGTVGVLRGDQGICWCGGPAFLDGCVCSAPLAELWVGVTYGLYNMPAWERRTRVE
 LEVDSELVYVGLTGTGSDTHSLSFLVRICHGFL"
 exon
 10891..11025
 /gene="At4g08830"
 /number=1
 11026..11243
 /gene="At4g08830"
 /number=1
 11244..11393
 /gene="At4g08830"
 /number=2
 11394..11429
 /gene="At4g08830"
 /number=2
 11430..12120

/gene="At4g08830"
 /number=3
 12121..12360
 /gene="At4g08830"
 /number=3
 12361..13089
 /gene="At4g08830"
 /number=4
 13090..13161
 /gene="At4g08830"
 /number=4
 13162..13585
 /gene="At4g08830"
 /number=5
 13586..13633
 /gene="At4g08830"
 /number=5
 13634..14348
 /gene="At4g08830"
 /number=6
 16002..16053
 18436..18437
 /gene="At4g08840"
 join(18436..19424,19870..20032,20124..20287,20492..20750)
 /gene="At4g08840"
 /note="similarity to probable repeat at AA82..92"
 -Saccharomyces cerevisiae, PIR2_53401;
 Contains Prokaryotic membrane lipoprotein lipid attachment
 site AA82..92"
 /codon_start=1
 /product="putative protein"
 /protein_id="CAB78009_1"
 /db_xref="GI:7267527"
 /translation="MDFGFPGDLRQROSFTL3PNG..LTSSVNGHEFSGDRTNPF
 FLMNRKLDITSLMDGVDMLCONL..AHSSPSYGYOGRESSING
 EASSMMSRGTGVDYHTRFEQDLRVAASPHEGESIMSYSVYDQSISGSDYHHRFLAQEASNP
 NPRECEENMSSLNNDYMLELHENQVIRDFSTVPKQS..LAPEERILPFSAMGGSR
 EVDGCSACKMKRNKEODSLDASMDSYGSYTMACQLC..-LORFVDIGNFYDVMIFK
 EVINNVIELGTDPRGNYLJOKLJTEVCNEQRQICLJLIF..SKPGLAVKISINNYGTRVY
 QKLIETWTTKEQISLVSXKSAVLPFGLSLERELNCVNVJ..CJKEFSPDNKFLIEATK
 FC..LIEATVPHGCVLQRCVSYSGEQHLQVLCV..LAQDPFNYLVOYIIEK
 VGGYNVFLPLRGNVYKLAQTKFSGHVVVKCLRYYPE..SGCIVNELSVLNGCYLQDPP
 YANVYIQCALSITKV.S"
 18436..19424
 /gene="At4g08840"
 /number=1
 19425..19869
 /gene="At4g08840"
 /number=1
 19870..20032
 /gene="At4g08840"
 /number=2
 20033..20123
 /gene="At4g08840"
 /number=2
 20124..20287
 /gene="At4g08840"
 /number=3
 20288..20491
 /gene="At4g08840"
 /number=3
 20492..20750
 /gene="At4g08840"
 /number=4
 23831..27580
 /gene="At4g08850"
 complement(join(23831..24120,24737..27580))
 CDS
 /gene="At4g08850"
 complement(join(23831..24120,24737..27580))
 /gene="At4g08850"
 /note="Ext GB:Fl4413 is reverse and marks the end of the gene
 similarity to receptor protein kinase-like protein -"

Arabidopsis thaliana	PIR2:T05898	contains Prokaryotic membrane lipoprotein lipid attachment site Aa12-22; Protein kinases signatures and profile AA763-784; Protein kinases signatures and profile AA883-895	contains EST gb:FL4412, Z17737, R84191, T22574, AA394763, AA395690, FL4413*
Query Match	29.4%	Score 499.8; DB 8; Length 179771;	
Best Local Similarity	90.6%	Pred. No. 1.7e-77;	
Matches	557;	Conservative 0; Mismatches 52; Indels 6; Gaps 2;	
y	1	TGGGTTTATGGATAAACATGACAATAATTATTATTCATGAGTTTATGGATA 60	
b	17646	TCGGTTTATGGATAAACATGACTAATATATTCATGAGTTTGTGGATAG 17587	
y	61	CATGACAAATTAAATATATCATGTTAAATAACATGTTAAATACATGCCATT 120	
b	17586	CATGACAAATTATTAATCATGTTAAATACATGTTAAATACATGCCATT 17527	
y	121	TTAAATCAGACATTGTTAAATCAATCTAATCTTATATCAACGACATTGC 180	
b	17526	TTAAATCAGACATTGTTAAATCAATCTAATCTTATATCAACGACATTGC 17467	
y	181	GGAAAATTAGGTTAAAAGGAAAATAAGAATCAGAGTAGAGATTCTATGGAAA 240	
b	17466	GGAAAATTAGGTTAAAAGGAAAATAAGAATCAGAGTAGAGATTCTATGGAAA 17407	
y	241	AGAAAGAGGAACATGTTAGGTGAAACAAAATAAGAGATAATGAGATATTATGAGAC 300	
b	17406	ATAAGAGGAAACATGTTGGTGACAATAAACAGATAATGAGATATTATGAGAC 17347	
y	301	GTTGTGAATTATT --- TAGGAGGGAGGAGAAATAGAAGAGAAAGACTCATGGCTA 355	
b	17346	GTTGTGAAGGATATTCTAGGAGGGAGGAGAAATAAGAATAGAAGAGAAAGACTCCATAGCTA 17287	
y	356	GTGAACTGAGAGATGGATGTGTTAAAGATGAAGAGAGAAAGAGACTCATGGCTA 415	
b	17286	GTGAACTGAGAGATGGATGTGTTAAAGATGAAGAGAGAAAGAGAAAGACTCCATAGCTA 17227	
y	416	AAGCTCTGTAAGAGAGATGAAAGAAACAAAAGGAGAGAAAGGAAAGGCTAA 475	
b	17226	AAGCTCTGPAAGAACATGAAATAAAAGAGAAAGGAAAGGCTAA 17167	
y	476	ATAGACTAACTATGGCCAAATTTCTGTAGCCGCAAAATACTATTGGCCAAGGTATT 535	
b	17166	ATAGACTAA-TATGGCCAAATTTCTGTAGCCGCAAAATACTATTGGCCAAGGTATT 17108	
y	536	TTGTGTATTTGAGCTAAAGTTATTCCTACATACTCTAAATAATAGCCGT 595	
b	17107	TTGTGTATTTGAGCTAAAGTTATTCCTACATACTCTAAATAATAGCCGT 17048	
y	596	ACCAATTTCACCA 610	
b	17047	AATACGTGCAA 17033	
RESULT	8		
DEFINITION	Arabidopsis thaliana BAC F9H3, from chromosome IV near 18.8 cm,		
ACCESSION	AF071527		
VERSION	AF071527.1		
KEYWORDS	HTG.		
ORGANISM	Arabidopsis thaliana		
AUTHORS	Huang, E.N., Parnell, L.D., de la Bastide, M., Schutz, K., Habermann, K., Dediha, N.N., and McCombe, W.R.		

/note="function=terminal_flanking_direct_repeat; cognate B
of flanking direct repeat of LTR transposon from position 5145 to
5149"
/rpt_type="flanking complement"(7885..7904)
/note="function="putative_microsatellite"
/rpt_type="tandem /rpt_unit="AT
complement"(7905..8818)
/note="function="unclassified; similar to MXI10, GenBank
accession number AB005248"
/rpt_type="dispersed
gene
/gene="T5iL23..3"
/note="encodes hypothetical protein; identical to F9H3.17,
GenBank accession number AF071527; gene model last edited
on 5 Jan 99"
/evidence="not_experimental
join(942..963,9745..9847,10690..10923)
/gene="T5iL23..3"
/note="identical to F9H3.17, genBank accession number
AC005560; similar to F21B7.5, GenBank accession number
AC005560; functional catalog ID="99"
/codon_start=1
/evidence="not_experimental
/product="hypothetical protein"
/protein_id="ADD15310..1"
/db_xref="GI:4263041"
/translation="MVKLRKRIGGLVRLRAFGALALAIVMITSRERASFLAISLEA
KYTDMAFKFVIANAVSVYSELVLFLESILWKFYVTLVYMTMLTSUSA
VAQVGKGNANAGWLPCIGGVPKFCDITGALIAGFVALVLYLVLSSHAVVDPF
LQKS"
12953..18635
/gene="T5iL23..4"
/note="encodes putative glucan synthase component;
identical to F9H3.18, GenBank accession number AF071527;
gene model last edited on 5 Jan 99"
/evidence="not_experimental
join(12953..14823,14984..17006,17187..18635)
/gene="T5iL23..4"
/note="similar to 1,3-beta glucan synthase; identical to
F9H3.18, GenBank accession number AF071527; functional
catalog ID=01..05..99; functional catalog ID=09..01"
/codon_start=1
/evidence="not_experimental
/product="putative glucan synthase component"
/protein_id="ADD15311..1"
/db_xref="GI:4263042"
/translation="MSLRHTTPPQTGRPLAAEAVGIEEPVNLIPIVNNLLADHPSLR
FPEVRAAAALKTVKTDLDRPQPYWQRSHDLDLMLAFGGFDQDNVRQREHMVLHLA
NAQMRLSPPDNIDSLSAVRERRKLKANYSSWC SYLKKENIWSIDRNDSRREL
AFLTGVLILLGEAANLRFAPCICNTYHMASELINKIILEDCLDENTGQYLPBLSGN
NEFKERGKVSGKTGEVERTEFEYLRSFELRLWMLALEFOQALIVANEKPKDTSVPR
QLWNALKARDQVRLILWVUTWSMRGLAVLDAQSOFVLSRETKRIFRMMLKVI
AAWVIVTAFTVLYNWQKRQDRSNAATKTYQFLVAGAFLPELALFLKIPW
MRNFLEETNMKIEFFALTWWFOQKFEVGRLEREGIYDNKYSTEWIFLATKEFESYFL
QVKPMIKSKLLWNLKDWDYHQEFYDNRFSTYALWMPVFLVLMQDQWAIYSS
IYGAVVGLDFHGLDFKLESNOVEANKTALIWELWILAFREDEEVLSREVELELPKNSWDY
LKLRYGFGRFKLESNOVEANKTALIWELWILAFREDEEVLSREVELELPKNSWDY
TWRNPCEFLCNEALLSQARELLDAPIPKLNHICKNAYRCWAWEADSLSKHL
SIKVDTETEISIITVFFQIINQSQEQSETKTFRDVLLKPIETLQKLVGLVDAEETD
SGRVNVNQSLYEATROFIEKTTTEQLSNEGTIPRPQSKULLFQNLRLDASNE
SYQRVRLHILTSRDSMSMSPVNLEARLTAFNSMNMPHQVKMAMSVLT
LRDRLWASYRGQOLARTGRMMYYRAKMKLAFLDSASEMDIREGAELGSVRNLQG
ELGGQSDGFSENDSSRSASSVSTLKGHENYKFTIVACOIGYSOKAKKE
PQAEEELLYLMKQNNEALRIAVYDEVPAFGRDTSVLYRQHOLEKEVETFRVKLPGP

Query Match 29.28; Score 497; DB 8; Length 116448;
Best Local Similarity 92.5%; Pred. No. 5.6e-77;
Matches 545; Conservative 0; Mismatches 40; Indels 4; Gaps 2;

repeat_region	Qy 1 TGGGTATTGATGAAATACATGACAAATTATTTCATGAGTTTATTGGATAG 60 Db 8481 TGATTTTATGGATACTGCAAAATTATCATGAATTACATGAATTGTTGGTAG 8422
repeat_region	Qy 61 CATGACAATATAATATATCATGTTATAACATGTTTGTCITTAATAATCATGCATT 120 Db 8421 CATGACAATATAATATCATGTTATAACATGTTTGTCITTAATAATCATGCATT 8362
repeat_region	Qy 121 TTAAATTCAGACATTTAACATTAATCTTATATCAAACGACATTCATGAC 180 Db 8361 TTAAATTCAGACATTTAACATTAATCTTATATCAAACGACATTCATGAC 8302
gene	Qy 181 GGAAATTCAGGAAAAAGAGATAAGAGATAAGAGATAAGAGATACTGATGGAA 240 Db 8301 GARGAATTCAGGAAAAGAGAAACAGAAGATAAGAGATACTGATGGAA 8242
CDS	Qy 241 AGAAAGAGAAACATGTTAGGTGAAACAATAAGAGATGATGATAATTATGAGAG 300 Db 8241 AGAAAGAGAAACATGTTAGGTGAAACAATAAGAGATACTGATGGAGAG 8182
gene	Qy 301 GTGTTGAACTTATTAGGAGGGAGAGAAATGAAAGAAAGATGGTGAAGCTA 360 Db 8181 GTGTTGAAAG-TATTCTAGGAGGGAGAAAGTAAAGAGAAATGATGGTGAAG 8123
CDS	Qy 361 TCTGAAGAAGATGATGTTAAAGATGAAAGAGAAAGACTCAAGCTAAGTGC 420 Db 8122 TTGAAAGAGATGATGTTAAAGATGAAAGAGAAAGACTCAAGCTAAGTGT 8063
gene	Qy 421 TCCTAAAGAGATGAAAGAAACAAAGAAAGGAAGAAAGAAAGCTAAATAGA 480 Db 8062 TCTGAAGAAGACATGAAAGAAAGAAAGAAAGAAATAATAGA 8006
RESULT 10	QY 481 CTAACATTGCCCCAAATTTCTGTAGGCCAAATACTATTGTCAGGTATTCTGT 540 LOCUS ATCHRV9
	DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment PLN 16-MAR-2000
	VERSION AL161497.2 GI:2720670
	KEYWORDS
	ORGANISM Arabidopsis thaliana
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicots; Rosidae; eurosids II; Brassicales; Arabidopsis.
REFERENCE	1 (bases 1 to 53060)
AUTHORS Spiegl,L.A., Huang,E.N., Nascimento,L.U., de la Bastide,M., Vil,D.M., Preston,R.R., Matero,A., Shah,R., O'Shaughnessy,A., Rodriguez,M., Shekher,M., Schutz,K., See,L.H., Svabý,I., Bertmann,K., Dediha,N.N., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.	
JOURNAL	Unpublished
REFERENCE	2 (bases 21279 to 137726)
AUTHORS	Zhong,J., Ma,P., Parnell,L.D., Chen,C.N., Chen,E.Y., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	3 (bases 137729 to 137730)
AUTHORS	Lamar,B., Stoeneking,T., Stumpf,F.J., Mewes,H.W., Lemcke,K. and Mayer,K.F.X.
JOURNAL	Unpublished
REFERENCE	4 (bases 1 to 159629)
AUTHORS	EU Arabidopsis sequencing project.
TITLE	Direct Submission
JOURNAL	Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:

lencke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de Project
 Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge
 Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK,
 E-mail: michael.bevan@bbsrc.ac.uk
 Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>
 this fragment has an overlap with ATCHRIV8 at the 5' end and an
 overlap with ATCHRIV10 at the 3' end.
FEATURES
source
 /organism="Arabidopsis thaliana"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="4"
 6741. 9588
 /gene="AT4g03480"
 complement(join(6741..7064..7161..7708..7783.. .8101..
 8409.. .886..9283.. .9343..9445.. .9588))
 CDS
 /gene="AT4g03480"
 complement(join(6741..7064..7161.. .7708..7783.. .8101..
 8409.. .886..9283.. .9343..9445.. .9588))
 /note="contains similarity to mammalian ankyrins
 similarity to similar to A. thaliana hypothetical protein
 F2B7_8, GenBank accession number AC02550
 Contains ARF/GRP-binding site motif A (P-loop) AA297..304
 contains EST gb:H37081, T420924
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB77833_1"
 /db_xref="GI:7270671"
 /translation="MVRSHLRATPTEASPLKPASSPSNCRRLSLSPVRLYFIRDDDF
 KVQDIFKTSYQNSLFLPCDLSKQGKQVPMPEIFAMRAGNI
 ELIEKLKSYETPNAKLSGDGSVLSVLAASCHLAEVKNLTTECPCLLLENSKTOIP
 LHYAARGRSAVVYLIVASLFLSRVPEDDRNLVYLKDGDPLHAKLDLHE
 KAEPERIKRQLSLHIMWRSRQIISDASTRQMAQASFILANKDGTSP
 LYLAVEAGNVSILVRAMINPGNKLOGISTLASOLPERKSLLHAALKAKNDVLAVIL
 NDDPSLVNEDEGRCTCLSVGASMGYTKICKLDRSTSKYECDKDGFSTIHMAYEK
 GHQKVKTEILKRCPDSSKELVNKGOKNMLHIAAKFASTETLUSKLNQDGURDINLHQDQY
 DVQGNAPIHLATINWRCTVTDKIAAFASTETLUSKLNQDGURDIAEUNLQDQY
 KRDHINANLLIVTTLVATV
 exon
 complement(join(6741..7064..))
 /gene="AT4g03480"
 /number=1
 complement(join(7065..7160..))
 /gene="AT4g03480"
 /number=1
 complement(join(7709..7782..))
 /gene="AT4g03480"
 /number=2
 complement(join(7783..8101..))
 /gene="AT4g03480"
 /number=3
 complement(join(8102..8408..))
 /gene="AT4g03480"
 /number=3
 complement(join(8409..8866..))
 /gene="AT4g03480"
 /number=4
 complement(join(8867..9282..))
 /gene="AT4g03480"
 /number=4
 complement(join(9283..9343..))
 /gene="AT4g03480"
 /number=5
 complement(join(9344..9444..))
 /gene="AT4g03480"
 /number=5

exon
 complement(9445..9588)
 /gene="AT4g03480"
 /number=6
 10335..13501
 /gene="AT4g03480"
 complement(join(10335..10555..10983.. .11078..11590.. .12133..
 12345..12500..12627.. .12772..12871.. .13501))
 /gene="AT4g03480"
 complement(join(10335..10525..10983.. .11078..11590.. .12133..
 12345..12500..12627.. .12772..12871.. .13501))
 /note="contains similarity to mammalian ankyrins
 similarity to similar to A. thaliana hypothetical protein
 F21B7_8, GenBank accession number AC02560"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CAB77834_1"
 /db_xref="GI:7270672"
 /translation="MGKQEFKVMENEIPVLDQVTFOGNTILHAAIYGHHDHVRR
 ILAYELNLRLRNKREGLNQNFPVSFSHYKOTLVERRNYKGDLAHVAAAGHKLIVGLI
 DCLRLPQDTMVIGSQDQVIGNFRVSINDGNTALHISLGKHHVSVSQLYEDRSTP
 CFLLDKEDPSPLMMAEAYSVLVEHMRGLDASFGVSVLCAVKSNDLITAVAE
 SDSDEYEVSEDEGTPLATAASGYDVOHMLTRFASTGQAVNLYVAKSNGNARAV
 GYLRLKSDYKRLINEQDTEGNTIPLHASSNSHPKVILWMAAGTAPRYHLRAD
 IPGIUTTDEDILKTHDRINTLTVLAVATMAGAAGSVPLGINSTEFKSNVXHSE
 ESAPHAFAVICNSTAVYTAVTAVTALIGTQDADLKCMITFKFIVPLIGFSIAMSIAF
 VAGLVLVGHYHWWIAFLASGGYLMLAIIILIPASDFTTLSRSLSNLSIONMSKE
 DVDSYVSGHAYDGGKKTSQL"
 complement(10335..10525)
 /gene="AT4g03490"
 /number=1
 complement(10526..10982)
 /gene="AT4g03490"
 /number=1
 complement(10983..11078)
 /gene="AT4g03490"
 /number=2
 complement(11079..11589)
 /gene="AT4g03490"
 /number=2
 complement(11590..12133)
 /gene="AT4g03490"
 /number=3
 complement(12134..12344)
 /gene="AT4g03490"
 /number=3
 complement(12345..12500)
 /gene="AT4g03490"
 /number=4
 complement(12627..12772)
 /gene="AT4g03490"
 /number=5
 complement(12773..12870)
 /gene="AT4g03490"
 /number=6
 13514..15185
 /note="marker mi233"
 14170..17288
 /gene="AT4g03500"
 join(14170..14204..15103.. .15848..15939.. .16477..16650.. .17288)
 /gene="AT4g03500"
 /note="hypothetical protein with some similarity to
 ankyrin
 similarity to"
 /codon_start=1
 /product="hypothetical protein"

/protein_id="CAB77835.1"
 /db_xref="IGI:7270673"
 /translation="MKNESETSEELTSKIGIFSCGVFSMGRDPOVILDRPVVAENHQIPIR
 NVRSPLSDJLSTDETSKPDKIMAVARAKENLRSNSYLSVAPLFLVNDRG
 NTILHAASSGHVSLLRVLIQKCPLLISNMGEVALHLAAARGHLDVWNULDFIN
 DISCNLPLVPAKRYFAKNNQDPAHVAKGKIEVVASVLSVSAKSUSPFLVANIDFESN
 LYLAEGHTSLWTTMCHGTELESSLVGRSIVHAALKANRKDLDALLSKDSLNL
 RDEGRFLSFGASIGYQGFSYLPDKNRVVKYFLSCCKDNKKNLINQDVNGNTPLH
 KHCPEAJELDRDGSNLLAAYKGKLYKWIKEVFLSCKDNKKNLINQDVNGNTPLH
 LATINWHPKVUSMFTWDHVYLDRKRNLYCFTALDVAEENDSSTVHORLTWALINA
 GAPKSTSPTEENLSFKKDDGGKTDKRVNTMLIVATLVMTAGFLPGYNDSPP
 HLGMAVLAKRTAFOVFLYCDTLLAMYSSTLITVIALWAGDLSITLKAFNIALPFLGL
 ALTSMSIAFMAGTYAVSHSLPLLGYEVLGIGIFFLVLLLPVYSPYAHQAQLLR
 IYYPFLKUAGDNKN1DVVTASDE"

exon
 14170..14204
 /gene="At4g03500"
 /number=1
 14205..15102
 /gene="At4g03500"
 /number=1
 15103..15848
 /gene="At4g03500"
 /number=2
 15849..15938
 /gene="At4g03500"

intron
 29 29.3% Score 497; DB 8; Length 159629;
 Best Local Similarity 92.5%; Pred. No. 5.3e-7;
 Matches 545; Conservative 0; Mismatches 40; Indels 4; Gaps 2;

Qy 1 TGGGTTTTATTGGATAACATGACAATAATTGTTAACATGCTTAATAATATGCCATT 120
 Db 29758 CATGACAATAATTGTTAACATGACAATAATTGTTAACATGCCATT 29639
 Qy 61 CATGACAATAATTGTTAACATGACAATAATTGTTAACATGCCATT 120
 Db 29698 CATGACAATAATTGTTAACATGACAATAATTGTTAACATGCCATT 29639
 Qy 121 TAAATTCAGACATTGTTAACATCAAACTCATCTATAATTCATGGATAG 29699
 Db 29638 TAAATTCAGACATTGTTAACATCAAACTCATCTATAATTCATGGATAG 29579
 Qy 181 GGAAATTCAAGTAAAAGGAAATAAAAGTAGAGATATGGAAAA 240
 Db 29578 GAAGAATTCAAGCAGAAAAGAACAAAGAACATAAGTAGAGATTCTATGGAAA 29519
 Qy 241 AGAAAAGAGAACATGTAGCTGAACAAATAAGAGATATGATGATATTATGGAG 300
 Db 29518 AGAAAAGAGAACATGTGGTGACAAATAAGAGATATGATGATAATTATGGAG 29459
 Qy 301 GTGGTGAAAGATTATTTAGAGGGAGAGAAATAAGAAANTGACATGGCAA 360
 Db 29458 GTGGTGAAAG-TATCTAGAGGGAGAAAGTAAAGAAATAATGACATGGCAA 29400
 Qy 361 TCTGAGAAAGATGAATTGTTAACATGAGAGAAAGACTCCATGGCTAAAGTC 420
 Db 29399 TTGTGAGAAAGATGAATTGTTAACATGAGAGAAAGACTCATGGCTAAAGTT 29340
 Qy 421 TCGTAAGAGATGAAAAAGAAACAAAGAGAGAAAGCTAAATAGAA 480
 Db 29339 TCGTAAGAGACATGAAAGAAAGAA--AAAGAAAGAAATAAAATAGA 29283
 Qy 481 CTAACATTCACAAATTCTGTAGCCGACAAACTATCTGTGCTAAAGGTATTGTG 540
 Db 29282 CTAATATGCCAAATTCTGTAGCCGACAAACTATCTGTGCTAAAGGTATTGTG 29223
 Qy 541 TATCTTGTGAAAGTCAAAGTTTCTCTACATATACTCTAAATAATA 589
 Db 29222 TATCTTGTGAAAGTCAAAGTTTCTGTAGCCGACAAACTATCTGTGCTAAATATA 29174
 RESULT 11
 ATH250185

LOCUS ATH250185 369 bp mRNA thaliana mRNA for NIMIN-2 protein (nimin-2 gene).
 DEFINITION PLN 15-AUG-2001
 ACCESSION AJ250185
 VERSION AJ250185.1 GI:12057155
 KEYWORDS nimin-2 gene; NIMIN-2 protein.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; eudicots II; Brassicales; Brassicaceae; Arabidopsis; Rosidae; eurosids II; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 369)
 AUTHORS Weigel,R., Bauscher,C., Pfitzner,U.M.
 TITLE NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from Arabidopsis that interact with NPRI/NTH1, a key regulator of systemic acquired resistance in plants
 JOURNAL Plant Mol. Biol. 46 (2), 143-160 (2001)
 MEDLINE 21335045
 PUBLMED 11442055
 REFERENCE 2 (bases 1 to 369)
 AUTHORS Pfitzner,U.M.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1999) Pfitzner U.M., German Virology, Institute of Genetics, Emil Wolff Str. 14, 70599 Stuttgart, GERMANY
 COMMENT Related sequence: AB023001.
 FEATURES Location/Qualifiers
 1..369
 /organism="Arabidopsis thaliana"
 /cultivar="Col-0"
 /db_xref="Taxon:3702"
 /tissue_type="green vegetative tissue"
 /dev_stage="3-wk-old"
 /gene="nimin-2"
 /product="NIMIN-2 protein"
 /protein_id="CAC19845.1"
 /db_xref="IGI:12057155.1"
 /db_xref="SPTRMBL:9LUAA3";
 /translation="MNNSLKEKEERVEEDNCSDGNRKPSSTEVVRTVTEEVDEFKILRRVHATRIVKYGNGVAGELPLSKRKRSLQRLNDCNGVRGEGDETEINRGLQGLGLLNLCPEDSVSLL"
 BASE COUNT 113 a 53 c 121 g 82 t
 ORIGIN 1
 Query Match 21.7% Score 369; DB 8; Length 369;
 Best Local Similarity 100.0%; Pred. No. 2.5e-54;
 Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 1 ATGAAACACTCTTGAAAGAAGAGACGCCATAGAAATCTGACGGT 922
 Qy 863 ATGAAACACTCTTGAAAGAAGAGACGCCATAGAAATCTGACGGT 922
 Db 1 ATGAAACACTCTTGAAAGAAGAGACGCCATAGAAATCTGACGGT 60
 Qy 923 AACAGAGGGAAACCGTGACCGGAAGTGTGGACCGTAACGGGAAGGGTGGAG 982
 Db 61 AACAGAGGGAAACCGTGACCGGAAGTGTGGACCGTAACGGGAAGGGTGGAG 120
 Qy 983 TTTCATCAAGATATTAGGAGATACAGTGGCAGACGAACTGTTGGACGGTACGGAGAGGGTGGAG 1042
 Db 121 TTTCATCAAGATATTAGGAGATACAGTGGCAGACGAACTGTTGGACGGTACGGAGAGGGTGGAG 180
 Qy 1043 GGTGTCCTGAGGGAGGTACCGCTCTAAGAGAAAGGGACCTGAGATCTGGGTG 1102
 Db 181 GGTGTCCTGAGGGAGGTACCGCTCTAAGAGAAAGGGACCTGAGATCTGGGTG 240
 Qy 1103 AGAAACTCGTGGATGTTGGATCTGAGGAAATTCGATTAATCGGTC 1162
 Db 241 AGAAACTCGTGGATGTTGGATCTGAGGAAATTCGATTAATCGGTC 300
 Qy 1163 GGTTACAGGTTGGATCTGAGGAAACAGACACUGGAAACAGACACUGGTTAGTTA 1222

RESULT 14
 AC115878/C LOCUS AC115878 171763 bp DNA linear HTG 14-JUN-2002
 DEFINITION Mus musculus clone RP24-35859, WORKING DRAFT SEQUENCE, 16 ordered
 pieces.

ACCESSION AC115878.2 TITLE HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS SOURCE house mouse.

ORGANISM Mus musculus

Eukaryote; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.

1 (bases 1 to 171763) AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Unpublished

2 (bases 1 to 171763) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gade,D., Galagan,J., Gardyni,N.,
 Ginde,S., Gord,S., Goyette,S., Hagan,J., Grand-Pierre,N.,
 Hagos,B., Horcon,L., Huime,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihowa,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,J.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Filey,R., Rose,C.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Youshikawa,T., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

JOURNAL (bases 1 to 171763)

REFERENCE Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihowa,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,J.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Filey,R., Rose,C.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL

REFERENCE AUTHORS

AC115878/C LOCUS AC115878 171763 bp DNA linear HTG 14-JUN-2002
 DEFINITION Mus musculus clone RP24-35859, WORKING DRAFT SEQUENCE, 16 ordered
 pieces.

ACCESSION AC115878.2 TITLE HTGS PHASE2; HTGS DRAFT; HTGS FULLTOP.
 KEYWORDS SOURCE house mouse.

ORGANISM Mus musculus

Eukaryote; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.

1 (bases 1 to 171763) AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Unpublished

2 (bases 1 to 171763) AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gade,D., Galagan,J., Gardyni,N.,
 Ginde,S., Gord,S., Goyette,S., Hagan,J., Grand-Pierre,N.,
 Hagos,B., Horcon,L., Huime,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,
 Mihowa,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,J.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Filey,R., Rose,C.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

JOURNAL (bases 1 to 171763)

REFERENCE AUTHORS

Result No.	Query Score	Match Length	DB ID	Description
1	230	13..5	230..17	BH866477 SALK_1013 BH866474 SALK_1013 BH866479 SALK_1013
c	2	203..8	12..0	531..17 BH695479
c	3	201..2	11..8	206..17 BH866474
c	4	185	10..9	237..17 BH814587
c	5	175..2	10..3	196..17 BH866500
c	6	173..6	10..2	267..17 BH854168 SALK_0788

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

8

RESULT 1
BH866477

DEFINITION SALK_101380 Arabidopsis thaliana TDNA insertion lines Arabidopsis

ACCESSION BH866477

VERSION 1 GI:22102375

KEYWORDS thale cress, Arabidopsis thaliana

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytina; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 230)
Alonso, J. M., Leisse, T. J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C. J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J. R.

TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome

JOURNAL Unpublished (2001)

COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)

The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.

FEATURES	Class: TDNA tagged.	Location/Qualifiers
Source		
1. .230		/organism="Arabidopsis thaliana" /strain="Columbia 0" /db_xref="Exon;3702"; /clone_id="SAK_101380"; /clone_lib="Arabidopsis thaliana TDNA insertion lines" /note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tDNA_protocols.html "
BASE COUNT	153 a	genomic DNA inserted into pHO51 using BstXI linkers"
ORIGIN	119 c	179 t
Query Match	13.5%	Score 230; DB 17; Length 230;
Best Local Similarity	100.0%	Pred. No. 1.1e-28; Gaps 0;
Matches	230;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	914	TGAGCTGTAAACAGGGAAACCGTGTGACGGAAAGTTGTCAGGGTAACGGGAAAG 973
Db	1	TGAGCTGTAAACAGGGAAACCGTGTGACGGAAAGTTGTCAGGGTAACGGGAAAG 60
QY	974	GTCATGAGTTTCAGATATTACCGAGACTACAGTGGCACACGAACGGTGTGCAA 1033
Db	61	GTCATGAGTTTCAGATATTACCGAGACTACAGTGGCACACGAACGGTGTGCAA 120
QY	1034	GTTAACGCCTGTGGAGGTACCGGTCAAGAGGAAACGGAGTCAGAT 1093
Db	121	GTTAACGCCTGTGGAGGTACCGGTCAAGAGGAAACGGAGTCAGAT 180
QY	1094	CCTGGGTGAGAAACCTGGTGAATTAAACGCCGTTGAGATACCGTCTAAGAGGAAACGGAGAAATT 1143
Db	181	CCTGGGTGAGAAACCTGGTGAATTAAACGCCGTTGAGATACCGTCTAAGAGGAAACGGAGAAATT 230
RESULT	2	BH866474
LOCUS	531 bp	DNA linear GSS 05-FEB-2002
DEFINITION	BOHWW78R BO_2_3_RB Brassica oleracea genomic clone BOHWW78, DNA sequence.	
ACCESION	BH866479	
VERSION	GI:18769095	
KEYWORDS	GSS.	
SOURCE	Brassica oleracea.	
ORGANISM	Brassica oleracea	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.		
REFERENCE	1 (bases 1 to 531)	
AUTHORS	Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.	
TITLE	Whole genome shotgun sequencing of Brassica oleracea	
JOURNAL	Unpublished (2001)	
COMMENT	Other GSSs: BOHWW78TF Contact: Chris Town TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtowen.tigr.org	
Class: sheared ends.	DNA is from a doubled haploid provided by Tom Osborn.	
Seq primer: TR	Seq primer: TR	
FEATURES	Location/Qualifiers	
Source	1. .531	
	/organism="Brassica oleracea" /strain="PO1000DH3" /db_xref="Taxon;3712"; /clone="BoHWW78"; /clone_id="Bo_2_3_KB"; /note="Vector: pHO51; Site_1: BstXI; 2-3 kb sheared	
BASE COUNT	153 a	
ORIGIN	119 c	
Query Match	12.0%	
Best Local Similarity	72.3%	
Matches	297;	
Conservative	0; Mismatches 102;	
Indels	12;	
Gaps	2;	
QY	950	GTTGGACGGTAAACGGAGAGGAGGTGGAGGTGGTGTGAGATATTAGGAGAGTAC 1009
Db	531	GTACGPACGGTGACCGAGAAAGGTGGTGTGAGATATTAGGAGACTAAC 472
QY	1010	GTGGGACACAGAACGGTTGCGAAAGTTAACGGGCGTGTGCTGAGGGAGTACCGCT 1069
Db	471	GTGGGACACGGATGTGGCAGAGTTAACGGGTTGACCTGAGTACCT 412
QY	1070	AAGAAGGAAACGGTCAAGAATCTGGTGTGAGAAACTCGTGTGGATTGAACTGGGT 1129
Db	411	AAGAAGGAAACGGATTAGGTTCTGGCTTGAGGCTCATGATACTAACGAGT 352
QY	1130	CGAGACGGAGAATTGGATGAGATAATCGGGTTACAGGGTTGGGTGGGATCTG 1189
Db	351	CAAGATGGAGATCAGATGAGATAAAATCGGGTGGPAGGGATTTGGPAGCT 292
QY	1190	AACTGTAACCGGAAACGACAGCGTTAGTTTACGTTGACTCTCATGT 1249
Db	291	AACTGTAACCGGAAACGGAGACTA-----AAATGTAATCTCTCTCTCATAA 243
QY	1250	TTTCCCTCTTCAATAATCAATTTCATAATTTCATAATTCAAAAATATG 1308
Db	242	CTAAAATATGTTCCAGTATCTATTCATAACATAACAAATAACAAATAAC 183
QY	1309	TAAAAGAAGATTAAACAGTTACCAATTTCAGAATCTCAGTTAT 1359
Db	182	TCTCTAAAACGTTACTAAATAATTATCCAAAGTCTGCTCTTGTAT 132
RESULT	3	
LOCUS	BH866474	
DEFINITION	BH866474 206 bp DNA linear GSS 05-AUG-2002	
ACCESSION	BH866474	
VERSION	GI:22102372	
KEYWORDS	GSS.	
SOURCE	Arabidopsis thaliana	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 206)	
AUTHORS	Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gaedrinah,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prendis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.	
TITLE	A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGnAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel: 858 453 4100 x1752 Fax: 858 558 6379 Email: ecker@salk.edu	
This is single pass sequence recovered from the left border of TDNA.	This is single pass sequence recovered from the left border of TDNA.	
FEATURES	Location/Qualifiers	
Source	1. .531	
	/organism="Brassica oleracea" /strain="PO1000DH3" /db_xref="Taxon;3712"; /clone="BoHWW78"; /clone_id="Bo_2_3_KB"; /note="Vector: pHO51; Site_1: BstXI; 2-3 kb sheared	
BASE COUNT	153 a	
ORIGIN	119 c	
Query Match	12.0%	
Best Local Similarity	72.3%	
Matches	297;	
Conservative	0; Mismatches 102;	
Indels	12;	
Gaps	2;	
QY	950	GTTGGACGGTAAACGGAGAGGAGGTGGAGGTGGTGTGAGATATTAGGAGAGTAC 1009
Db	531	GTACGPACGGTGACCGAGAAAGGTGGTGTGAGATATTAGGAGACTAAC 472
QY	1010	GTGGGACACAGAACGGTTGCGAAAGTTAACGGGCGTGTGCTGAGGGAGTACCGCT 1069
Db	471	GTGGGACACGGATGTGGCAGAGTTAACGGGTTGACCTGAGTACCT 412
QY	1070	AAGAAGGAAACGGTCAAGAATCTGGTGTGAGAAACTCGTGTGGATTGAACTGGGT 1129
Db	411	AAGAAGGAAACGGATTAGGTTCTGGCTTGAGGCTCATGATACTAACGAGT 352
QY	1130	CGAGACGGAGAATTGGATGAGATAATCGGGTTACAGGGTTGGGTGGGATCTG 1189
Db	351	CAAGATGGAGATCAGATGAGATAAAATCGGGTGGPAGGGATTTAGCTG 292
QY	1190	AACTGTAACCGGAAACGACAGCGTTAGTTTACGTTGACTCTCATGT 1249
Db	291	AACTGTAACCGGAAACGGAGACTA-----AAATGTAATCTCTCTCTCATAA 243
QY	1250	TTTCCCTCTTCAATAATCAATTTCATAATTTCATAATTCAAAAATATG 1308
Db	242	CTAAAATATGTTCCAGTATCTATTCATAACATAACAAATAACAAATAAC 183
QY	1309	TAAAAGAAGATTAAACAGTTACCAATTTCAGAATCTCAGTTAT 1359
Db	182	TCTCTAAAACGTTACTAAATAATTATCCAAAGTCTGCTCTTGTAT 132

/clone_lib="Arabidopsis thaliana TDNA insertion lines"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at <http://signal.salk.edu/tDNA-protocols.html>"

BASE COUNT	82 a	be found at http://signal.salk.edu/tDNA-protocols.html "
ORIGIN	38 c	63 g
	53 t	1 others
Query Match	11.88;	Score 201.2; DB 17; Length 206;
Best Local Similarity	98.54;	Pred. No. 6.8e-24;
Matches	203;	Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	938	TGACCGAAGTGTGGACCGTAACGGACGAAGAGGTGGATGAGTTTCAGATATA 997
Db	1	TGACGGACTTGATGGTACGGTAGGAAAGGGATGAGTTTCAGATATA 60
Qy	998	CGGAGAGTACACGTGGCACGACGAGTTGCGAAAAGTTPAACCGCGGTGTGTGAGGA 1057
Db	61	CGGAGAGTACAGTGGCACGACGAGTTGCGAAAAGTTAACCGCGGTGTGTGAGGA 120
Qy	1058	GAGTTACCGCTAAAGAGGAAACGGAGTCAAGATCTGGGTGAGAAACTGTTGAT 1117
Db	121	GAGTTACCGCTAAAGAGGAAACGGAGTCAAGATCTGGGTGAGAAACTGTTGAT 180
Qy	1118	TGAAAGGGTTGAGACGGAGAAATT 114.3
Db	181	TGAAAGGGTTGAGACGGAGAAATT 206

RESULT 4

LOCUS	BH814587	237 bp DNA linear GSS 02-MAY-2002
DEFINITION	SALK_066674	Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_066674, DNA sequence.
VERSION	BH814587	GI:20394318
KEYWORDS		
SOURCE	Arabidopsis thaliana	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis thalei cross.	
REFERENCE	Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.	
TITLE	A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGNAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
FEATURES	source	
source	1..196	/organism="Arabidopsis thaliana" /strain="Columbia 0"
Class:	TDNA tagged.	/db_xref="taxon:3702" /clone_id="SALK_101406"
Class:	TDNA tagged.	/note="PCR was performed on Arabidopsis thaliana TDNA insertion lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tDNA-protocols.html "

RESULT 5

LOCUS	BH865500	196 bp DNA linear GSS 05-AUG-2002
DEFINITION	SALK_101406	Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_101406, DNA sequence.
ACCESSION	BH865500	GI:22102398
KEYWORDS	GSS	
SOURCE	Arabidopsis thaliana	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis thalei cross.	
REFERENCE	Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J., and Ecker,J.R.	
TITLE	A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGNAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA	
FEATURES	source	
source	1..196	/organism="Arabidopsis thaliana" /strain="Columbia 0"
Class:	TDNA tagged.	/db_xref="taxon:3702" /clone_id="SALK_066674"
Class:	TDNA tagged.	/note="PCR was performed on Arabidopsis thaliana TDNA insertion lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tDNA-protocols.html "

page 9

OY 485 CTTATGCCAAAAA 496
DB 248 PAAAAAAAAAA 237

Search completed: November 29, 2002, 00:31:28
Job time : 2058 secs

ALIGNMENTS

RESULT 1

US-09-983-965-2109/C
Sequence 2109, Application US/09983965
; Patent No. US2002013716A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21102971C
; CURRENT APPLICATION NUMBER: US/09/983_965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2109
; LENGTH: 529
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (335)
; OTHER INFORMATION: Other information: Clone ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

result	Query				ID	Descript
	No.	Score	Match	Length	DB	
c	1	73.8	4.3	529	10	US-09-983-965-2109
c	2	72	4.2	446	10	US-09-960-352-3400
c	3	71.2	4.2	152331	9	US-10-095-407-16
c	4	71.2	4.2	176373	9	US-10-095-407-17
c	5	68	4.0	292	10	US-09-867-701-9583
c	6	68	4.0	516	10	US-09-960-352-5785
c	7	67.8	4.0	180216	10	US-09-835-232-6
c	8	66.4	3.9	174424	10	US-09-967-768A-314
c	9	66.4	3.9	299	10	US-09-867-701-9596
c	10	66.2	3.9	307	10	US-09-867-701-9416
c	11	65.6	3.9	42999	9	US-09-799-462B-17
c	12	65.6	3.9	42999	9	US-10-125-767-17
c	13	65.2	3.8	451	10	US-09-960-352-10262
c	14	63.8	3.8	174424	10	US-09-967-768A-314
c	15	63.4	3.7	428	10	US-09-960-352-573
c	16	63	3.7	299	10	US-09-867-701-9477
c	17	63	3.7	341	10	US-09-960-352-1202
c	18	62.6	3.7	152331	9	US-10-095-407-16
c	19	62.4	3.7	1721	9	US-09-938-922-3573

Qy 227 ATTCTATGGAAAAAGAGAAACATGTAGGTGACAAAATAAGAGATATGTATGAT 286
 Db 354 AAAACAAAACAaaaaaaaANAAAACAAAAAaAAAaAAAaAAAaAAAaAAA 295

RESULT 3
 US-10-095-407-16/c
 ; Sequence 16, Application US/10095407
 ; Patent No. US20020164330A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 ; FILE REFERENCE: 09404/052001
 ; CURRENT APPLICATION NUMBER: US/10-095,407
 ; CURRENT FILING DATE: 2002-03-11
 ; PRIOR APPLICATION NUMBER: US 60/091,650
 ; PRIORITY NUMBER: US 60/091,650
 ; PRIORITY FILING DATE: 1998-07-02
 ; PRIORITY APPLICATION NUMBER: US 60/054,646
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 16
 ; LENGTH: 152331

RESULT 2
 US-09-960-352-3400/C
 ; Sequence 3400, Application US/09960352
 ; Patent No. US20020137139A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nengbing
 ; APPLICANT: Byatt, John C.
 ; ORGANISM: Bos taurus
 ; OTHER INFORMATION: Clone ID: 15-LIB3058-052-Q1-K1:D11
 US-09-960-352-3400

Query Match 4.2%; Score 72; DB 10; Length 446;
 Best Local Similarity 48.5%; Pred. No: 1.4e-05;
 Matches 198; Conservative 0; Mismatches 210; Indels 0; Gaps 0;

Qy 108 AATACATGCCTTAACATCAGACATTGTGTTAACATCAATCTTATATCA 167
 Db 409 AAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAA 350

Qy 168 CAACGACATGACCGAACATTCACTGTAAGAAATAAGAGATAGAGA 227
 Db 349 AAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATA 290

Qy 228 TTCTATGGAAAAAGAGAAACATGTAGGTGACAAAATAAGAGATATGTATGATA 287
 Db 289 AAATAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATA 230

Qy 288 TATTATGAGGGTGGTGAAGATTATTTAGGAGGGAGAATAGAAAAAGAAA 347
 Db 229 ATATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 170

Qy 348 ATGACATGGTGAATCTGAGAATGAAATGTATAGATGAGAAGAGACTC 407
 Db 169 AAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAAaAAA 110

Qy 408 CATGGCTAACAGTCGTAAGAGATGAAAGAAACAAAAGAGGAAGAGAA 467
 Db 109 AAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATA 50

Qy 468 AGGCTAAAATGACTAACTATGGCAAATTCTGTAGCGGACAAAATA 515

RESULT 4
 US-10-095-407-17/c
 ; Sequence 17, Application US/10095407
 ; Patent No. US20020164330A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pan, Yang
 ; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
 ; FILE REFERENCE: 09404/052001
 ; CURRENT APPLICATION NUMBER: US/10-095,407
 ; CURRENT FILING DATE: 2002-03-11
 ; PRIOR APPLICATION NUMBER: US 60/091,650
 ; PRIORITY NUMBER: US 60/054,646
 ; PRIORITY FILING DATE: 1998-07-02
 ; PRIORITY APPLICATION NUMBER: US 60/091,650
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO: 17

Page 4

```

    /> Patent No. US20020098489A1
    /> GENERAL INFORMATION:
    /> APPLICANT: Leder, Phillip
    /> APPLICANT: Leader, Benjamin
    /> TITLE OF INVENTION: FORMIN-2 NUCLEIC ACIDS AND POLYPEPTIDES
    /> FILE REFERENCE: 00383/052002
    /> CURRENT APPLICATION NUMBER: US/09/835,232
    /> CURRENT FILING DATE: 2001-04-12
    /> PRIOR APPLICATION NUMBER: US 60/196,811
    /> PRIOR FILING DATE: 2000-04-13
    /> NUMBER OF SEQ ID NOS: 22
    /> SOFTWARE: FastSEQ for Windows Version 4.0
    /> SEQ ID NO: 6
    /> LENGTH: 180216
    /> TYPE: DNA
    /> ORGANISM: Homo sapiens
    /> FEATURE: misc_feature
    /> NAME/KEY: misc_feature
    /> LOCATION: (1)...(180216)
    /> OTHER INFORMATION: n = A,T,C or G
    /> US-09-835-232-6

        Query Match          4.0%;  Score 67 ;  DB 10;  Length 1
        Best Local Similarity 49.4%;  Pred. No. 0.0046;
        Matches 205;  Conservative 0;  Mismatches 207;  Indels

    Qy   167 ACAACGACATTGAGCGAAAATTCAAGTAAAGAGAAATAAACGATGAGAG
    |   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   85794 ACAAGACAGAGACAGACAGAGACGGACGGACGGAGACAAACAGAGATACAGAG
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   227 ATTTCTTGAAAGAAAGAGAGAACATGTAGTGTGAAACAAATAAAAGAT
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   85854 AGAGAGAGACAGAGACAGACAGAGACAGAGAGAGAGAGACAGAGACAGAAACAG
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   287 ATATTTTATGAGAGCTGGTGAAGATTATTTTGGAGAGGGAGAAATAG
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   85914 ACAGATAAAAGAGAAAGAGACAGAGACAGAGACAGAGAGAGAGAAAG
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   347 ATAGACATGTGAATCTGTGAGAAGATGAAATTGTTAAGATGAGAGAA
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   85974 AAATACAGAGAGACAGAGACAGAGACAGAGAA--GATACGAAAGAGATGAC
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   407 CCATGGCTAAAGTCCTGTAAGAAGATGAAAGAAACAAAGAGGGAGAA
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   86031 AGAAAGAGAGAGAACAGGGAAAGAAAAGAGAGAGAAAGAGAGAGAA
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   467 AAGCTTAAGATGACTAACTATGCCAAATTCTGTGACCCGAGAAATACTA
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   86091 AGGAGAGAGAGAGACTGAGACTGAGTTGTCCTCAATTGGGTAAACAGAA
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Qy   527 AAGGTATTTGTATCTTGAAGTAAAGTTATTCTCATACATACTC
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |
    Db   86151 GAATTAACACCATTTTAACTACAGACCTTCATTTAATTTCTCT
    |   | | | | | | | | | | | | | | | | | | | | | | | | | | |

    RESULT 8
    US-09-967-768A-314
    /> Sequence 314, Application US/09967768A
    /> Patient No. US2002015087A1
    /> GENERAL INFORMATION:
    /> APPLICANT: AUGUSTUS, Meena
    /> TITLE OF INVENTION: Cancer Gene Determination and Therapeutic
    /> TITLE OF INVENTION: Sets
    /> FILE REFERENCE: 689290-72
    /> CURRENT APPLICATION NUMBER: US/09/967,768A
    /> CURRENT FILING DATE: 2001-09-28
    /> PRIOR APPLICATION NUMBER: US/60/236,109
    /> PRIOR FILING DATE: 2000-09-28
    /> PRIOR APPLICATION NUMBER: US/60/236,034
    /> PRIOR FILING DATE: 2000-09-28
    /> PRIOR APPLICATION NUMBER: US/60/236,111
    /> PRIOR FILING DATE: 2000-09-28

```


Page 6

